Genetic Architecture of Linear Localized Scleroderma

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Rebecca Geraldine Higgins

aus

Irland

Promotionskommission

Prof. Dr. med. Dr. sc. nat. Alexander Navarini (Leitung der Dissertation, Vorsitz)

Prof. Dr. med. Anita Rauch

Prof. Dr. Burkhard Becher

Zürich, 2018

Table of Contents

Chapter 1

The Genetic Architecture of Linear Localised Scleroderma

Table of abbreviations

ADEN	Acantholytic dyskeratotic epidermal nevi	**LoSSI**	Localized scleroderma severity index
AI	Autoimmune	**LS**	Localized scleroderma, morphea
AS	Apert syndrome	**MAF**	Minor allele frequency
BAM	Binary alignment map	**MHC**	Major histocompatibility complex
BL	Blaschko's lines	**MMF**	Mycophenolate mofetil
CADD	Combined annotation dependent depletion	**MTX**	Methotrexate
CARRA	Childhood Arthritis and Rheumatology Alliance	**NGS**	Next-generation sequencing
CD	Cluster of differentiation	**NSAID**	Nonsteroidal anti-inflammatory drug
CGH	Comparative genomic hybridization	**PBMC**	Peripheral blood mononuclear cell
CNS	Central nervous system	**PCR**	Polymerase chain reaction
CNV	Copy number variation	**PGA-A**	Physician global assessments of activity
CyS	Cyclosporin A	**PGA-D**	Physician global assessments damage
CS	Corticosteriods	**PRS**	Parry Romberg syndrome
ddNTPs	Dideoxynucleoside triphosphate	**PUVA**	Psoralen and ultraviolet A
dNTPs	Deoxynucleoside triphosphate	**SAM**	Sequence alignment map
DP	Read depth	**SF**	Snap frozen
DS	Darier's disease	**SNP**	Single nucleotide polymorphism
ECDS	En coup de sabre	**SNV**	Single nucleotide variation
EPACTS	Efficient and Parallelizable Association Container Toolbox	**SS**	Sanger sequencing
ESP_MAF	Exome sequencing project minor allele frequency	**SSc**	Systemic sclerosis
FFPE	Formalin-fixed paraffin embedded	**SSC**	Squamous cell carcinoma
GATK	Genome analysis toolkit	**TG_MAF**	Thousand genome project minor allele frequency
GRC	Genome reference consortium	**UCSC**	University of California, Santa Cruz
HLA	Human leukocyte antigen	**UDG**	Uracil DNA glycosylase
IGV	Integrative genomic viewer	**VCF**	Variant called format
LLS	Linear localized scleroderma	**VEP**	Variant effect predictor
LOH	Loss of heterozygosity	**WES**	Whole exome sequencing
LoSCAT	Localized scleroderma assessment tool	**WGS**	Whole genome sequencing
LoSDI	Localized scleroderma damaging index		

Abstract (English)

Linear localized scleroderma (LLS) is a rare connective tissue disorder (<2.7/100'000 per year) characterized by chronic inflammation and massive accumulation of collagen. This results in hardening and thickening of the lesion leading to the affected areas to cave in from atrophy. The sharply delimited and linear lesions can affect patients anywhere on the skin and mostly on the face in the clinical subtype's *en coup de sabre* and Parry-Romberg syndrome. This leads to disfigurement and impaired quality of life. The disease affects mostly children and treatment options are limited and most often unsatisfactory.

Very little is understood about the condition in terms of genetic and clinical aetiology. There is evidence that LLS might be based on genetic alterations in affected tissues. Blaschko's lines are the patterns of cell migration and proliferation during embryological development. Multiple skin conditions have been shown to follow Blaschko's lines including LLS. Several of these diseases are caused by genetic factors such as a *de novo* somatic mutation producing cutaneous mosaicism. Here, we test the hypothesis that LLS is caused by a somatic genetic protein-coding mutation.

The whole exome of blood and affected skin taken from 19 confirmed LLS patients was sequenced. Oligonucleotides complementary to the protein-coding DNA regions were used for library preparation. Sequence analysis was performed with GATKv3.5. To further validate results, 4 patients with similar lesions were chosen for deep sequencing (DP of >300). However, somatic analysis of these samples revealed no potentially causative mutation. CGH was performed on 3 patients to find large scale chromosomal aberrations too large to be detected by WES. Samples were compared with 3000 control karyotypes. Germline analysis of samples against 80 controls also did not reveal any driving factor. HLA signals were

determined between the 19 cases and 60 controls, no allotype was significantly associated with LLS. A limitation of this project is the lack of intronic, intergenic and epigenetic data. Taken together, the in-depth analysis performed in this work allowed us to disprove the hypothesis that a somatic exonic event causes LLS.

Abstract (German)

Linear lokalisierte (zirkumskripte) Sklerodermie (engl.: linear localized scleroderma = LLS) ist eine seltene Bindegewebserkrankung (Neuerkrankungen <2.7 pro 100.000 Personer im Jahr), welche sich in chronischer Entzündung und massiver Ablagerung von Kollagen in der Haut manifestiert. Durch diese Anreicherungen von Kollagen kommt es zu verhärteten Hautläsionen, welche sich durch Atrophie wieder abflachen. Die scharf abgegrenzten und linearen Läsionen können überall auf der Haut auftreten, jedoch speziell bei den klinischen Subtypen *en coup de sabre* und *Parry-Romberg Syndrom* finden sich diese im Bereich des Gesichtes. Dies führt regelmässig zur Entstellung und beeinträchtigter Lebensqualität. Die Krankheit tritt am häufigsten bei Kindern auf und die Behandlungsmöglichkeiten sind limitiert und unbefriedigend.

Es ist nur wenig bekannt über die genetischen und klinischen Ursachen dieser seltenen Krankheit. Es gibt aber Hinweise, dass lineare lokalisierte Sklerodermie auf genetischen Mutationen in den betroffenen Bereichen basiert. Die sogenannten Blaschko-Linien beschreiben embryonale Wachstumslinien der Haut, entlang dieser während der Embryonalentwicklung Zellproliferation und Migration stattfindet. Multiple Hautveränderungen sind den Blaschko-Linien zugeordnet, einschliesslich linear lokalisierte Sklerodermie. Einiger dieser Erkrankungen werden durch genetische Faktoren ausgelöst, wie zum Beispiel *de novo* somatische Mutationen, welche ein kutanes Mosaik erzeugen. In dieser Arbeit wurde die Hypothese getestet, ob linear lokalisierte Sklerodermie durch somatische Mutationen in protein-kodierenden Genombereichen ausgelöst wird.

Aus Blut und betroffener Haut von 19 Patienten mit gesicherter linear lokalisierter Sklerodermie wurde DNA isoliert und alle darauf befindlichen protein-kodierenden Abschnitte (Exon) sequenziert (engl. WES=whole exome sequencing). Mit Hilfe von für protein-codierende Sequenzen komplementären Oligonukleotiden wurde aus der Patienten-DNA eine sogenannte «library» hergestellt. Diese besteht aus amplifizierten DNA Fragmenten, welche später mit bekannten DNA Fragmenten hybridisiert und sequenziert werden. Die Resultate der Sequenzierung wurden mit Hilfe der GATKv3.5 Software analysiert. Um die Resultate weiter zu validieren, wurde bei 4 phänotypisch sehr ähnlichen Patienten ein noch gründlicheres «deep sequencing» durchgeführt. Hierbei werden mindestens 300 amplifizierte DNA Fragmente des gleichen Typs sequenziert. Mit Hilfe der CGH Technologie (comperative genomic hybridization) können grosse chromosomale Veränderungen analysiert werden, welche durch die WES Methode nicht detektiert werden. Die DNA Proben von 3 Patienten wurden mit Hilfe von CGH analysiert und zu 3000 Karyotypen aus Kontrollgruppen verglichen. Die Keimbahnmutationen der Proben wurden auch mit 80 Kontrollproben verglichen. Schliesslich wurden auch die HLA Klasse I Allotypen bestimmt und die 19 Patienten mit der Normalpopulation verglichen. In allen diesen Untersuchungen zeigte sich keine krankheitsverursachende Mutation, chromosomale Aberration oder HLA-Signal . Eine Limitation dieses Projektes ist das Fehlen von intronischen, intergenetischen und epigenetischen Daten.

Zusammengefasst konnte mit den in dieser Arbeit durchgeführten intensiven und detaillierten genetischen Analysen die Hypothese widerlegt werden, dass somatische Mutationen für die Entstehung der linear lokalisierten Sklerodermie eine Rolle spielen.

Introduction

Linear Localized Scleroderma

1. Background

Linear localized scleroderma (LLS) is a rare disorder of the connective tissue. It is characterized by chronic inflammation and massive deposition of collagen and an increase in extracellular matrix production. This leads to hardening and thickening of the area, which eventually caves in from atrophy leaving the patients with an indurated linear lesion(s) (Weibel and Harper 2008, Fett and Werth 2011, Majewski, Schwartzentruber et al. 2011). The disease is not fatal but can lead to cosmetic disfigurement, especially if the lesion is located on the face. LLS is generally limited to cutaneous and subcutaneous tissues and is self-limiting in most cases (Weibel and Harper 2008). In some, more severe, cases there may be involvement of the muscle and bone underlying the lesion. The extracutaneous involvement can lead to side effects such as arthritis, epilepsy, arthralgia and uveitis (Christen-Zaech, Hakim et al. 2008, Fett and Werth 2011). The disease is not always self-limiting and can progress over the years causing irreversible structural deformities such as joint contractures and mental disabilities like epilepsy (Weibel and Harper 2008).

LLS is a subtype of localized scleroderma (LS). LS is more commonly known clinically as morphea. Morphea is the preferred term in the clinical setting to avoid patient confusion with systemic sclerosis (SSc) which could cause unnecessary confusion and concern. LLS is the subtype of morphea where plaques or lesions are in a linear form. The linear presentation of morphea is the most common of the five subtypes to be found in children. Pediatric morphea

is estimated to be of the linear subtype in 40-70% of cases (Kreuter, Krieg et al. 2009, Weibel, Laguda et al. 2011, Marsol 2013).

The average age of onset of LLS is roughly 6 years of age (Zulian, Athreya et al. 2005). However congenital presentation has been described in both morphea and LLS (Zulian, Vallongo et al. 2006). The estimated incidence of morphea is 1-2.7 per 100,000 individuals. There is as of yet no specific data available for the incidence of LLS. This could be due to the difficulty in diagnosis of morphea and the difficulty in differentiation of LLS from other subtypes of morphea. Women are estimated to be affected by morphea 2.1-6 times more often than men and specifically LLS has been found in two studies to be 2.2 times more likely in women than men (Orozco-Covarrubias, Guzmán-Meza et al. 2002, Zulian, Athreya et al. 2005, Tollefson and Witman 2007).

Previous studies have been ethnically diverse and despite the presence of morphea in all races, it has been shown to affect predominantly Caucasians (72.7-82%) (Amaral, Peres et al. 2013). In a study of 44 patients with LLS by Christen-Zaech *et al.* 2008, 37 were Caucasian, 5 were Hispanic, 1 Asian and 1 African American. The same ethnic distributions were also found in other subtypes of morphea (Christen-Zaech, Hakim et al. 2008). As of yet there is no explanation for the unequal distribution in ethnicities.

2. Classification

The first system of classification of the subtypes of morphea was described in Peterson *et al*. 1995. This study concluded there are 5 subtypes of morphea; Plaque (keloidal sclerosis, atrophoderma of Pasini and Pierini guttate, lichen sclerosis), Generalized (morphea lesions on two or more areas), Bullous, Deep (eosinophilic fasciitis, Pansclerotic morphea, morphea profunda, subcutaneous morphea) and Linear (ECDS, PRS, linear lesion of extremities) (Peterson, Nelson et al. 1995). This classification system over time has not been widely accepted, as many of the disorders such as eosinophilic fasciitis are separated as a distinct disease. In addition, this classification system does not take into account the 15% of cases where patients have more than one subtype of morphea (Fett and Werth 2011). 67% of morphea in children is reported to be LLS. The second most common subtype is plaque type morphea (26%), followed by generalized morphea (7%) and deep morphea (2%).

Another system of classification was described by Laxer and Zulian in 2006 and has now become the more widely accepted version. This new system of classification was necessary as some diseases in the old classification system were not actually morphea and did not account for the occurrence of more than one subtype of morphea in a patient. This classification system is the commonly used system in clinical analysis of morphea patients. This classification also describes 5 subtypes (Table 1) (Laxer and Zulian 2006, Fett and Werth 2011). These subtypes are circumscribed, generalized, pansclerotic, linear and mixed (Figure 1).

The circumscribed subtype is the most common type of morphea which has two subtypes, superficial and deep. The superficial subtype is limited to the cutaneous regions and the deep subtype can extend into muscle. Generalized morphea has widespread involvement and is

most often limited to the dermis. The mixed subtype is described in roughly 15% of morphea patients. This is when a patient has two morphea subtypes, most commonly one of them will be LLS. The most debilitating morphea subtype is pansclerotic. Its effects can extend from the dermis into the bone and result in chronic wounds. Morphea is generally not fatal however, fatal pansclerotic morphea has been described in a study of an 11-year-old patient with a history of pansclerotic morphea on her trunk, which over 2 years expanded to affect her entire body. It rapidly became ulcerative and eventually resulted in death (Hardy, Audouin-Pajot et al. 2016). Pansclerotic morphea has also been linked to squamous cell carcinoma (SCC) with 6.7% of pansclerotic morphea patients estimated to develop SCC (Fett and Werth 2011).

Subtype	Description
Circumscribed /Plaque morphea	Ovular areas of induration. Has two subtypes, superficial and deep. Superficial type is limited to the dermis and epidermis. Deep type extends into subcutaneous level and underlying bone.
Generalized morphea	This describes widespread lesions, ≥4 plaques on ≥2 sites of the body.
Linear localized scleroderma	Linear lesions on limbs and trunk. Also includes facial subtypes ECDS and PRS. Involves dermis, subcutaneous and rarely muscle and bone.
Pansclerotic morphea	Circumferential involvement of skin, subcutaneous, muscle and bone.
Mixed morphea	Involves more than one subtype to be present. Most commonly, the linear subtype is found alongside another subtype. It occurs in up to 15% of patients.

Table 1. Classification of morphea into 5 subtypes as proposed by Laxer and Zulian in 2006.

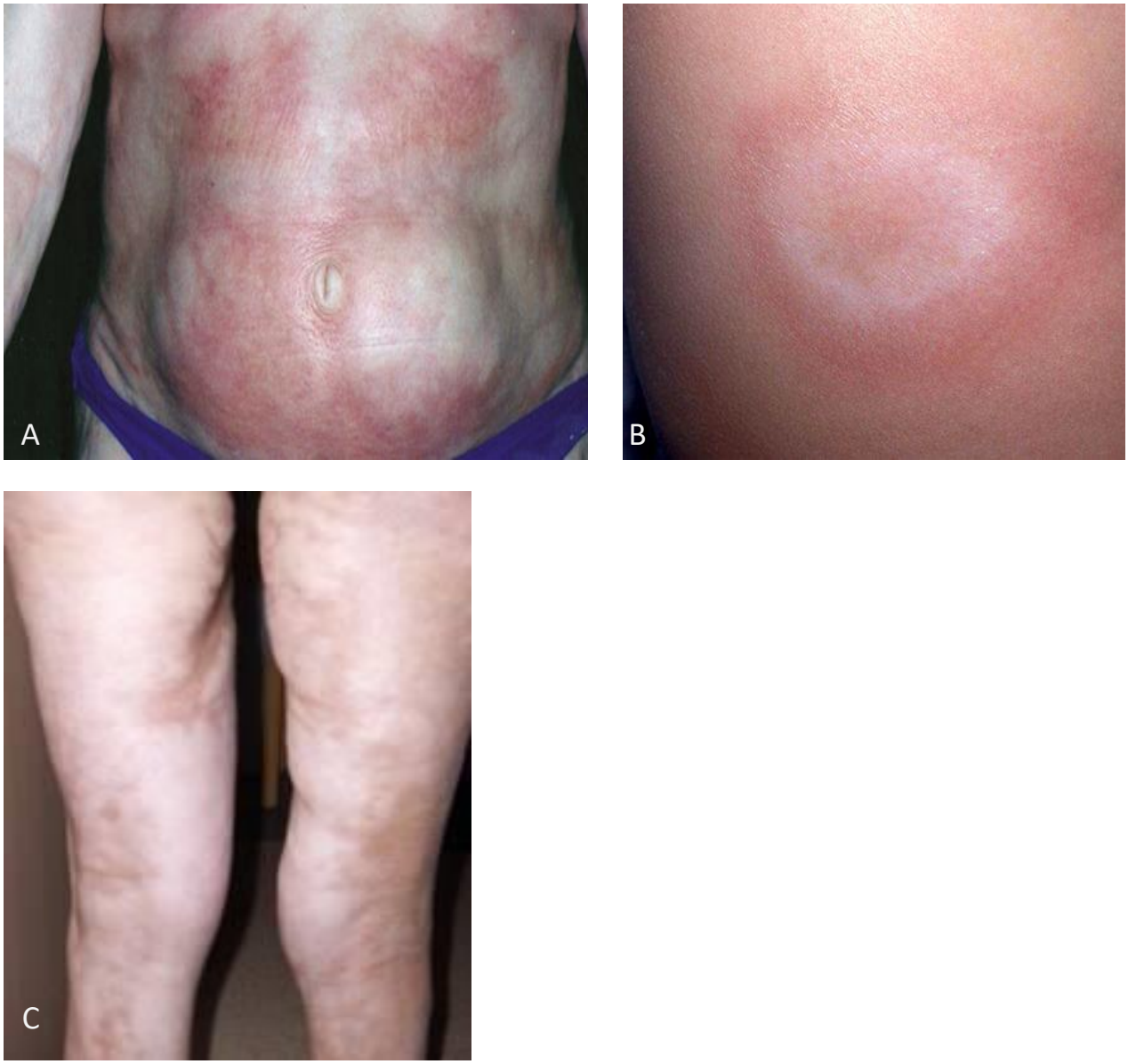


Figure 1. Subtypes of morphea other than LLS.
A) Generalized morphea (Neustadter, Samarin et al. 2009),
B) Circumscribed morphea superficial (www.cursoenarm.net) and
C) Pansclerotic morphea (Ana-Maria, Aura-Nicoleta et al. 2008)

3. Pathogenesis and triggers

Pathogenesis has been shown to be largely similar in LLS, morphea and SSc in terms of skin involvement. Early skin biopsies in all three diseases show damaged endothelial cells sometimes years before fibrosis. Vessels in the area lose their elasticity due to the outer layers of the vessels becoming rigid and fibrotic. A high concentration of mononuclear lymphocytic cellular infiltrates in dermal and subcutaneous regions have been described in morphea skin biopsies. This indicates inflammation as a precursor to fibrosis (Fleischmajer, Perlish et al. 1977). It has also been shown that fibroblasts of scleroderma skin produce more collagen than those of normal skin (Johnson and Ziff 1976, Rodnan, Lipinski et al. 1979).

Figure 2.
Hyperpigmentation from inflammation in early stage ECDS

The “classic evolution” of morphea begins with an erythematous patch that slowly becomes yellowish and white and features blue-violet erythema at the edge. Subsequently, post-inflammatory hyperpigmentation of underlying tissues ensues (Figure 2). This is generally considered the early-stage of the disease (Kreuter, Krieg et al. 2009). The progression of LLS can vary from patient to patient. Often the early inflammatory stage can be intensely erythematous, leading to a misdiagnosis of common port-wine stain (Weibel, Laguda et al. 2011).

In contrast, the inflammatory stage at the skin surface may be delayed with atrophy of underlying tissue occurring rapidly. Clinical features that can help in distinguishing LLS from other diseases in early stages, is the distribution along Blaschko’s lines, loss of hair on the scalp or eyebrow, skin fibrosis, atrophy of the skin and age of the patients. When lesions are found on the face it is also important to search for facial asymmetry from atrophy of underlying tissues (Figure 3) (Weibel, Laguda et al. 2011).

The late-stage of the disease develops when the affected area begins to harden and thicken (Kreuter, Krieg et al. 2009). This is caused by a massive increase of collagen in the area. This collagen is closely packed becoming dense and leads to fibrosis. The fibrotic regions become hard and discoloured and eventually atrophy will occur (Torok and Arkachaisri 2012). Atrophy in the area can lead to impairment of growth and limb function, especially in growing children. It has been shown that the skin that is affected in all subtypes of morphea can soften over time when the area has been inactive for a long time. However the other types of damage such as atrophy and dyspigmentation are permanent. In LLS, the contractures caused by muscle damage are permanent (Szramka-Pawlak, Dańczak-Pazdrowska et al. 2014).

Figure 3. Asymmetry from LLS lesion on the left paranasal area. Figure from our own clinical cohort.

What trigger causes this damage remains unknown as yet. Many events and environmental stimuli have been implicated. Mechanical damage appears to be the most cited culprit: A traumatic event to the area has been described as a precursor to disease manifestation in many studies (Yamanaka and Gibbs 1999). In a large-scale study by Zulian and Athreya et al. 2005 based on the clinical evaluation of 750 young patients with morphea, 489 patients were determined to have LLS. 100 patients were studied to unearth the preceding factors that possibly lead to onset of morphea. 69 of these patients had LLS. 51 patients reported the onset was due to mechanical factors (40 = trauma, 8 = Insect bites and 3 = vaccination). Of

the patients who reported trauma, 12 patients reported that it was due to infection, 3 reported it was a response to drug therapy and 3 reported it was due to previous psychological distress (Zulian, Athreya et al. 2005). Sunburn and lichen striatus that preceded the development of morphea have also been reported (Christen-Zaech, Hakim et al. 2008). Childhood head injury was implicated by 27% of respondents in an internet survey of 205 patients (Stone 2003).

Infections with *Borrelia burgdorferi*, commonly known as Lyme disease, have also been implicated in the disease progression. This assertion is an area of contention. One study concluded that infections with Lyme disease have late-stage skin involvement in the form of fibrotic nodules, which eventually progress to other skin diseases like morphea (Malane, Grant-Kels et al. 1991). This association between morphea and Lyme's disease has been widely questioned with some papers finding a significant relationship between *B. burgdorferi* antinuclear antibodies in the blood of morphea patients (Prinz, Kutasi et al. 2009), whilst others did not confirm any associations (Hoesly, Mertz et al. 1987, Kreuter, Krieg et al. 2009). *B. burgdorferi* survives in certain environments, so its geographical distribution has been well mapped out. The same can also be said for clinical incidence of Lyme disease. As of yet no work has been done to investigate the relationship between morphea incidence vs Lyme disease incidence from a geographic standpoint.

4. Complications and co-diagnosis of LLS and morphea

For the patient, aside from the skin involvement, LLS is usually asymptomatic with sometimes itchiness and mild pain reported. The largest problem for children is the gradual growth of healthy tissue that is not matched by the atrophied and hardened lesions. This effect can even lead to impaired bone growth in patients with lesions on arms and legs. Joint movement can also be restricted. ECDS and PRS can also result in seizures as the lesion extends towards the brain (Fett 2013).

In a partially retrospective study of paediatric morphea in 136 patients, non-cutaneous manifestations were found to be more common in the 44 LLS patients who had lesions on the body rather than the face. Several patients had musculoskeletal issues due to LLS. Arthralgia was diagnosed in 13 patients. Muscular cramps and scoliosis was diagnosed in four and two patients respectively. 10 patients with lesions on the face had CNS involvement which was predominantly migraines, seizures and stroke. Of the 44 patients with no facial lesions, only three had CNS involvement. These patients had seizures, headaches and dyspnea. Another three patients also had gastrointestinal problems such as reflux, dysphagia and general abdominal pain (Christen-Zaech, Hakim et al. 2008).

In this same study of the patients with ECDS (n=23) and PRS (n=3), none of the patients were diagnosed with musculoskeletal side effects and gastrointestinal problems. Of the ECDS patients, 10 were affected by headaches and seizures with one patient suffering a stroke. No CNS involvement was found in the PRS patients (Christen-Zaech, Hakim et al. 2008).

This study also investigated co-morbidity with other autoimmune diseases (AI) in LLS patients. The AI history of extended and immediate families were also included. In the patients with non-facial LLS, two patients were found to have juvenile rheumatoid arthritis. Vitiligo,

psoriasis, alopecia areata and Crohn's disease were also described in individual patients. In the family members, five had rheumatoid arthritis, two had systemic lupus, two had thyroiditis, two had systemic scleroderma and one had celiac disease. Family members of the three PRS had no history of AI. One PRS patient suffered from alopecia areata. One ECDS patient had juvenile dermatomyositis. In family members of ECDS patients, only three family members had AI such as autoimmune thyroiditis, psoriasis or type I diabetes.

Some studies have reported fatigue, myalgia, malaise and arthralgia appearing with morphea (Fett and Werth 2011). Rheumatologic, ophthalmologic and neurologic symptoms have also been described in many LLS patients (Amaral, Marques Neto et al. 2012).

LLS is also the most likely form of morphea to present with another subtype of morphea resulting in a diagnosis of the mixed subtype (Fett and Werth 2011). In a study of 750 patients with morphea 15% of children had the mixed subtype. LLS was found alongside circumscribed, superficial and deep forms. LLS lesions appeared before or at the same time as the other plaques in 64% of patients. The mixed form is not easily recognizable at the clinical level and confusion can result. The treatment for the mixed subtype is difficult as treatment options for different subtypes vary (Zulian, Athreya et al. 2005).

5. Diagnosis of LLS

Differential diagnosis of LLS from other subtypes is difficult in some patients. Some confusion can occur in patients with mixed diagnosis, however there are 4 factors that are helpful in separating LLS from the other subtypes:

1) Depth of the lesions: LLS has extracutaneous involvement, which is not common in generalized and superficial circumscribed morphea.
2) The linear nature of the lesions with sharply delimited borders are the most significant indicators of LLS.
3) The patient's age is also very helpful in diagnosis. LLS is the most common morphea subtype in children and its onset is roughly 6 years of age.
4) Adherence to BL is also an indicator of LLS.

The initial diagnosis of LLS or morphea can be difficult. There are the classical symptoms and onset as described previously but the agreed "classical evolution" (discussed in section 1.3) is most often not the way both diseases progress. There is significant variation in the clinical presentation, hence the rate of misdiagnosis is non-negligible. Factors involved in pathogenesis in morphea and LLS are important for diagnosis. Early stage yellowish-white erythematous patches with violet edges are helpful in diagnosis. After the inflammatory stage, hyper-/hypopigmented lesions are also indicative of morphea and LLS. More helpful in diagnosis is the hardening, loss of elasticity and thickening of the skin.

Here, some tools a techniques used in diagnosis of morphea and LLS are described.

5.1 Histology

Dermatopathology reveals the microscopic pattern of inflammation and structural pathology, which is crucial for the diagnosis of LLS. All subtypes of morphea are divided into two phases, early and late. In the early phase, there are often dense periadnexal and perivascular inflammatory in the dermis. Cellular infiltrates are also found, mainly lymphocytes, plasma cells and histiocytes. More rarely, eosinophilic granulocytes are present. Edema is sometimes found in the upper dermis. Thickened collagen fibres are found in the upper dermis. Late stage morphea is recognizable by the presence of sclerosis and atrophy of sweat glands. Also, the walls of blood vessels are thicker and there are signs of inflammation (Figure 4) (Kreuter, Krieg et al. 2009).

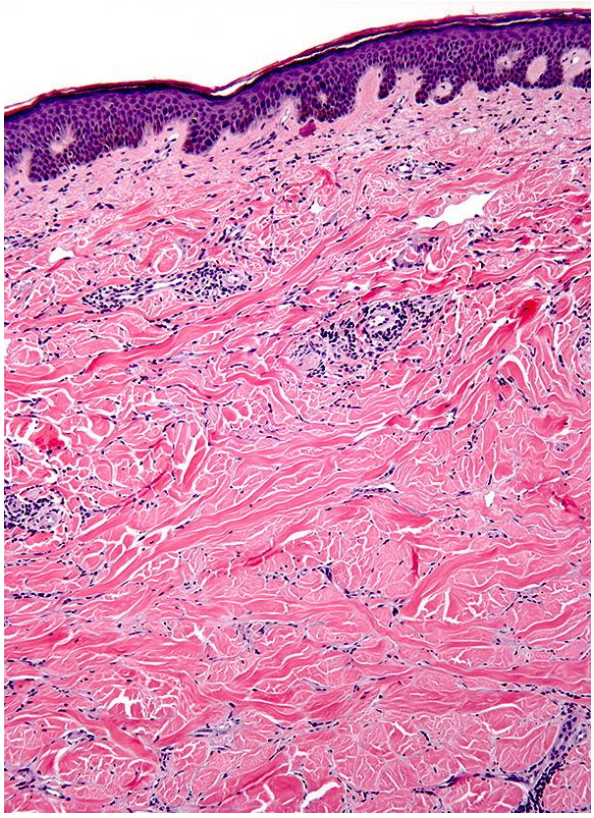


Figure 4. Histology of early morphea shows lymphocyte infiltrate and enlarged endothelial cells and collagen bundles. Hematoxylin-eosin stain; original magnification, x10^3 (Fett and Werth 2011).

5.2 Laboratory techniques

A myriad of laboratory tests can be utilized to aid the clinician in their diagnosis. Some diseases can be identified promptly as these test results can clearly implicate a certain disease. Therefore it would be attractive to find a serologic marker, cellular infiltrate or autoantibody that directly implicates morphea as the correct diagnosis. Furthermore, identifying a factor to differentiate LLS from other subtypes of morphea would be very interesting and potentially useful as well.

Laboratory tests performed on 474 LLS patients at the time of diagnosis showed some abnormalities, and discordance between other subtypes of morphea. The patients were tested for their eosinophil cell blood count, which was reported to be abnormal in 12% of LLS patients (57/474), this rate was higher in circumscribed superficial and generalized morphea (18.4% and 13.7% respectively) and much higher in the circumscribed deep form (62.5%) (Zulian, Athreya et al. 2005).

Erythrocyte sedimentation rate (ESR) was increased in 22.2% of patients (102/460) similar to that of deep morphea with an abnormal level in 25% of patients. In 249 tested patients with LLS, IgA, IgG and IgM was increased in 12.4%, 20.9% and 16.1% of patients respectively. These results are similar to plaque morphea (13.5%, 14.9% and 8.1%) and generalized morphea (11.8%, 11.8% and 9.5%). IgA and IgM in deep morphea are similar to other types of morphea (18.8% and 6.2%) however, IgG was much higher at 56%. Antinuclear antibodies (ANA) were tested in 446 LLS patients and found positive in 47.3% of patients. This is similar to other morphea subtypes ranging from 31.3%-43% positive ANA serologies (Zulian, Athreya et al. 2005).

Antihistone antibodies and ssDNA are usually positive in LLS as well (Sartori-Valinotti, Tollefson et al. 2013). These factors can be helpful in diagnosis but are largely limited as no statistical work has found a true correlation.

5.3 Measurement tools that are considered in assessment

Thermography

Thermography has been suggested as a method to measure disease activity. This uses infrared imagery to assess the surface temperature of the skin. It involves acclimatizing in a temperature-controlled room after which the images are taken using an infrared camera. The surface temperatures are contralaterally compared, focusing on affected and non-affected regions. It has been found that there is a significant but small increase of 0.29°C in lesions that are active.

This small change in temperature could be explained by inflammation and microcirculatory changes in the lesions. This change in temperature is interesting. The usefulness of this technique is limited and only relevant in active lesions and there have been false positives found in inactive lesions (Birdi, Shore et al. 1992, Garcia-Romero, Randhawa et al. 2017). This has been theorized to be due to the lack of subcutaneous tissue from the disease increasing thermal conductivity of the epidermis. Given that it is not possible to distinguish active and inactive lesions, using thermographic measurements may not be useful in diagnosis (Garcia-Romero, Randhawa et al. 2017).

Durometer

This is a tool, which measures and evaluates skin hardness of the affected skin relative to the contralateral unaffected area. Due to collagen build-up in the area there is a notable increase in hardness in affected areas. This tool depends on factors such as oedema, sex and age of patients, and location (Fett and Werth 2011). It is easy to use and has been shown to have low variability between patients. It unfortunately has low correlation with skin scores and is therefore not commonly used (Fett and Werth 2011).

Cutometer

A cutometer is a device to measure the elasticity and relaxation of skin. A vacuum pump is applied to an 8 mm diameter area of skin. Elasticity is assessed by measuring skin disruption followed by time for skin to return to baseline (relax) (Fett and Werth 2011). Elasticity of the skin is affected by the fibrosis and atrophy of LLS affected skin. This technique has been shown to be helpful in other skin diseases but has not been validated in any studies on LLS (Dobrev 2005, Kreuter, Krieg et al. 2016).

Ultrasound

Use of a 20 MHz ultrasound has been shown to have use in diagnosis of LLS. It can be used to visualize the top centimetres of the skin. Echogenicity is altered especially in the later stages of morphea. As edema and inflammation decrease echogenicity increase. Furthermore sclerosis increases as the disease progresses and this results in a corresponding increase in echogenicity. Use of a 20 MHz ultrasound shows a thickened and hypoechoic dermis that is

not present in normal skin. This method has been shown to be reproducible, it has recently been validated (Bendeck and Jacobe 2007, Kreuter, Krieg et al. 2009).

6. Data collection in LS and LLS

Decisions on treatment and disease management options can be very difficult for a practitioner especially if there is no standardized criteria for assessment and treatment. There are several reasons why no double-blind studies have been performed on any of the available morphea treatments. The main obstacle is that it is a rare disease, thus gathering sufficient volumes of data and information is difficult. Furthermore, in the case of rare diseases and especially non-fatal ones, pressure to gather data of good and comparable quality is not omnipresent.

The most widely used score for morphea and the score used in this study is the Localized Scleroderma Severity Index (LoSSI). This score describes the disease activity by 4 assessment criteria. Each criteria is scored in a scale of 0 to 3 based on least to most severity with the exception of the 4th criteria.

1) The surface area score. 14 cutaneous sites are assessed throughout the body. Each area is scored on a scale of 0 to 3 with how much of the surface area of the site is affected. A score of 0 at one site means there is no lesion. A score of 3 means over 2/3 of the site is covered by a lesion.
2) The erythema score. This scores the inflammation by severity focusing on the edges of the lesion and scoring from 0: normal to 3: dark red or marked erythema.

3) Skin thickness. This is assessed by palpation of the skin which is scored on a scale of 0 to 3. 0 is normal skin and 3 is marked increase in thickness and limited mobility to the skin.
4) Development of new lesions. If a new lesion appears within one month of the last consultation a score of 3 is given.

However, there are limitations of this score the most important of which is it doesn't take damage caused by the disease into consideration. There is a score describing damage called the Localized Scleroderma Damaging Index (LoSDI). This score describes the level of damage to the area. Damage is defined as irreversible or persistent change (Arkachaisri, Vilaiyuk et al. 2009). The three criteria are dermal atrophy (DAT), subcutaneous atrophy (SAT) and hyper-/hypopigmentation all of which are scored on a scale 0-3 (0 = normal skin, 3 = severe phenotype).

It can be concluded from literary research that patients with all subtypes of morphea have limited treatment options, the success of which vary drastically. The biggest issue facing doctors is the lack of standardized diagnosis and treatment protocols for patients. The Childhood Arthritis and Research Alliance (CARRA) are attempting to establish such a protocol, however it is proving slow and difficult. The best solution for this is to create standard and verified useful tools for collection of patient data.

Due to the phenotypic variability and subsequent difficulties / delays in diagnosis, thoroughly collected data cannot be compared inter-patient. Creation of a "gold standard analysis" for practitioners should alleviate uncertainty, workload and inaccuracy. A new skin scoring technique has been described called Localized Scleroderma Assessment Tool (LoSCAT). This tool combined the LoSSI and LoSDI. This method has shown some promise in data collection

hopefully leading to standardization and laying the groundwork for extensive testing of treatments (Arkachaisri, Vilaiyuk et al. 2009). LoSCAT attempts standardize the assessment of morphea patients between doctors in treatment centres worldwide. This will facilitate complete characterization of the rare disease and strengthen meta-analysis. This could lead to double-blind testing of drugs on patients and hopefully a more personalized and thus effective treatment approach in patients (Marsol 2013).

7. Subtypes of facial LLS

There are two subtypes of LLS that are found on the face and head, *en coup de sabre* (ECDS) and Parry-Romberg syndrome (PRS). Both disorders are often mistaken for each other at the clinical level due to the facial asymmetry that results from both and neurological issues that can also accompany both diseases (Orozco-Covarrubias, Guzmán-Meza et al. 2002). No diagnostic markers to completely differentiate the two diseases exist. However, cutaneous markers, histopathology and clinical examinations are very useful in the differential diagnosis. In spite of this, there are still many cases where differentiating the two has eluded even the most experienced histopathologists and clinicians. It is for this reason that it remains difficult to completely describe and understand the relationship of these two LLS facial subtypes.

7.1 En Coup de Sabre (ECDS)

ECDS is one of the subtypes of LLS located on the head. It is so called due to its likeness to a strike from a sword. EDCS lesions are hyperpigmented, sclerotic, shiny and hairless. Alopecia of the area that arises if the lesion is in or extends to the scalp. This is often the reason the patient first seeks medical attention (Figure 5) (Orozco-Covarrubias, Guzmán-Meza et al. 2002). Lesions are most often located on the frontoparietal scalp, median and paramedian forehead and sometimes extend into the scalp, maxillary area and side of the nose (Figure 6) (Tollefson and Witman 2007). This pattern has also been noted to follow the upper trigeminal nerve (Gambichler, Kreuter et al. 2001). In most cases, ECDS will not extend past the midline of the face but it has in some cases extended down to the upper lip or chin (Figure 7). It

appears unilaterally, however a few cases of bilateral involvement has been noted (Rai, Handa et al. 2000, Gambichler, Kreuter et al. 2001).

ECDS appears in most patients before 10 years of age (Orozco-Covarrubias, Guzmán-Meza et al. 2002). Congenital presentation of ECDS has been described, however, it is very rare (Zulian, Vallongo et al. 2006).

Figure 5. Alopecia in ECDS, extending from the forehead into the scalp. Figure from our clinical cohort.

Figure 6. ECDS A) lesion on the right forehead B) Second lesion of ECDS on the right temporal of the same patient. Some alopecia present. Lesion from skin biopsy C) Lesion on right temporal region of a patient extending to the eye of a second patient. All figures from our clinical cohort.

Figure 7. ECDS affecting the chin. Figure from our clinical cohort.

Some studies have shown severe headaches can act as a harbinger of the disease (Polcari, Moon et al. 2014). In 4 patients it was observed that severe headaches started 6 months to 3 years prior to cutaneous onset. In 3 of the 4 cases the headaches appeared localized to the area where the cutaneous manifestation developed. These four female patients had severe photophobia, dizziness and nausea during their headache spells. The frequency and duration of the headaches appear to have no relationship. In one patient, they were as frequent as 45-minute spells 5 times per day and once a month in another patient. The four patients had MRIs but intracranial irregularities in the form of signal intensity at many loci, ipsilateral the cutaneous findings, were found in only one patient. This patient had the most severe phenotype having suffered from unilateral transient weakness which lead to “stroke-like” symptoms, vomiting, migraines and photophobia up to 3 years before the first lesion manifested (Polcari, Moon et al. 2014).

7.2 Parry-Romberg Syndrome (PRS)

PRS is also known in the literature as Progressive Facial Hemiatrophy and more rarely Idiopathic Hemifacial Atrophy (Kreuter, Krieg et al. 2016). It is a unilateral, slow progressing and self-limiting degeneration of tissues beneath the skin of the face leading to hemifacial collapse of varying severities (Blaszczyk, Królicki et al. 2003). It affects subcutaneous tissues and there appears to be no cutaneous involvement such as hyperpigmentation or tissue hardening as found in ECDS (Orozco-Covarrubias, Guzmán-Meza et al. 2002).

Atrophy affects the underlying muscles and bones which results in one side of the face to cave in (Figure 8). It usually takes between 2-10 years for the disease to stop progressing. Some studies have described dormant PRS lesions to be triggered to accelerate. In the majority of these cases, the apparent trigger was pregnancy and childbirth (68%) with other triggers reported to be surgery and stress (Tolkachjov, Patel et al. 2015). Some women have reported the PRS lesion to worsen during pregnancy (Stone 2003). There is also reported to be involvement with the gums, trigeminal nerve, tongue and palate (Jun, Kim et al. 2011)

PRS onset is generally before 20 years however, it was estimated in a study of 54 patients that the average age of onset is 13.6 years with a range of 0.3 to 75 years (Tolkachjov, Patel et al. 2015). 66-80% of patients with PRS are female and age of onset is generally before 20 with the median age being 10 years old (Stone 2003, Tolkachjov, Patel et al. 2015).

Neurologic abnormalities, mainly epilepsy has been reported in 10% of PRS patients with some reporting seizures. Migraine and facial pain have also been reported. (Blaszczyk, Królicki et al. 2003, Stone 2003). PRS was originally described by Romberg as a trophoneurosis i.e. the malfunction of trophic nerves (Henoch and Romberg 1846). This theory was later tested by disruption of the superior cervical sympathetic ganglion on one side through surgery in animal

models (cats, dogs and rabbits) when they were 30 days old. Monthly examinations were performed in the following year. The investigators found the development of PRS-like symptoms in the animals namely alopecia, ulceration, ocular atrophy and hemifacial atrophy. This suggests a relationship between malfunction of the sympathetic nervous system (SNS) and PRS (Resende, Dal Pai et al. 1991, Blaszczyk, Królicki et al. 2003). This is interesting when you consider many neurologic side effects have also been found in the CNS.

Like ECDS, the distribution of the PRS atrophy is often observed along the three division lines of the trigeminal nerve, which is part of the CNS. The three divisions or branches are the ophthalmic nerve, maxillary nerve and mandibular nerve. The divisions and secondary branches span the majority of the face. The 3 divisions are joined and surrounded by SNS fibres (Go, Kim et al. 2001). Alternatively, the origin of PRS has been suggested to be linked to neurocranial malformations and slow viruses (Orozco-Covarrubias, Guzmán-Meza et al. 2002). In addition, vascular damage can be seen in PRS patients, possibly due to inflammation of vessels (Tolkachjov, Patel et al. 2015).

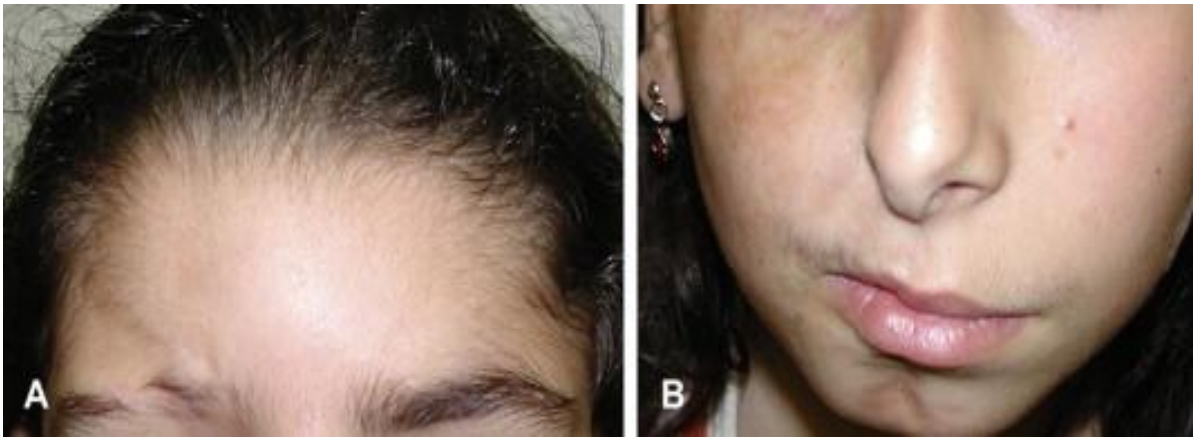


Figure 8. A) PRS on the right forehead focused on the paramedian region,
B) PRS affecting the cheek and lip of the same patient resulting in severe atrophy and asymmetry (El-Kehdy, Abbas et al. 2012).

PRS does not have cutaneous sclerotic involvement, therefore treatment options are better than in ECDS. Once the patient reaches maturity, reconstructive surgery can be sought, whereas in ECDS the patient requires active drug therapy (Orozco-Covarrubias, Guzmán-Meza et al. 2002).

In a small, retrospective study by Sommer *et al*. 2006 it was suggested that there could be two subtypes of PRS. Type 1 affects the subcutaneous tissues, primarily the cheek. Type 2 first affects the skin and will move on to the deeper tissues as it progresses. The second variant overlaps with ECDS. This accounts for the level of overlap and diagnostic confusion.

In the Sommer *et al*. 2006 study of 12 patients with PRS type 1 patients (n=5) all but one had tongue, lip, salivary gland and gingiva involvement, often in combination. No patients had involvement of the trunk or a co-diagnosis of ECDS. Neuroimaging of these patients was all normal however 2 patients had neurological symptoms (linguistic and motor delay and psychosis).

In type 2 PRS patients (n=7) two patients had tongue involvement, one had palsy (nerves III, IV, VI, VII) and one patient had a fixed pupil. ECDS was also diagnosed in 5 of these patients. One patient had deep morphea, one had LLS and one had plaque morphea. 2 of the 7 patients did not have neurologic symptoms. The other patients suffered from a range of neurologic symptoms from seizure (Jacksonian, *Grand mal*) to migraine and one had depression. Neuroimaging was not available for two patients and was found to be normal in three patients. One patient with no neurologic symptoms had normal neuroimaging whereas the other normal patient had no neuroimaging (Table 2) (Sommer, Gambichler et al. 2006).

Patient	Type	Onset (years)	Side	Morphea co-diagnosis	Intraoral and opthamological involvement	Neuologic symptoms	Neuroimaging	EEG
1	1	3	R	-	Tongue and lips	Jacksonian seizures and migraine	Frontoparietal subcortical hyperintense lesion in FLAIR and T1	Overlaying beta activity
2	1	3	R	Deep morphea	Tongue	Normal	Normal	Parietal sharp wave focus
3	1	7	L	ECDS	-	Grand mal and migraine	Normal	Normal
4	1	6	Bilateral	LLS and ECDS	Palsy of nerves III,IV, VI, VII	Grand mal, palsy of nerves	Frontoparietal subcortical hyperintense lesion in FLAIR and T2	Intermittent bilateral frontotemporal theta-delta-activity
5	1	27	R	ECDS	-	Depression	-	-
6	1	4	L	Plaque morphea and ECDS	Fixed pupil	Grand mal	Ventricular dialation, subcortical calcification	Normal
7	1	10	L	ECDS	-	Normal	-	-
8	2	3	R	-	Salivary glands, tongue and lip	Normal	Normal	Parietal sharp wave focus
9	2	2	R	-	Gingiva and tongue	Linguistic and motor delay	Normal	Temporal sharp wave focus, generalized delta-rhythm on sleep EEG
10	2	9	L	-	Upper lip and tongue	Psychosis	Normal	Normal
11	2	16	L	-		Normal	Normal	Normal
12	2	7	R	-	Tongue	Normal	Normal	Normal

Table 2. Table with clinical data of 12 patients with PRS separated into two subtypes based on cutaneous involvement. Adapted from (Sommer, Gambichler et al. 2006)

7.3 Neurologic symptoms in ECDS and PRS

In a study investigating the neuroabnormalities of ECDS and PRS patients they combined the retrospective in-house clinical data along with the data from reviews of available literature (Chiu, Vora et al. 2012). In respect to the literature review 73% of patients were reported to have abnormalities from MRI and CT scans, whereas in the in-house review, only 19% of all patients had an abnormality. The main abnormalities found were lesions on T2 sequences that reveal inflammation and oedema. In this study, MRIs were performed on patients that were asymptomatic and symptomatic for any neurological problem. Abnormalities were found in 22% of MRI scans of symptomatic patients' scans however abnormalities were also found in 17% of the asymptomatic patients. Also there are commonly bony or intracranial manifestations of the disease that are visible by MRI as well as scalp tissue abnormalities (Polcari, Moon et al. 2014).

This study has highlighted the importance of neuroimaging in all patients with ECDS and PRS regardless of neurologic involvement not only to help with treatment options but also to predict future problems.

7.4 Distinction between PRS and ECDS

PRS and ECDS are often thought of as overlapping disorders with some researchers of the thought that PRS is just a more severe and widespread form of ECDS. Conversely, some would argue they are completely separate disorders.

The most distinguishing clinical feature is the cutaneous involvement in ECDS that is not present in PRS, but this is not always sufficient to distinguish the two. Both entities show atrophy and thinning at the affected site as well as sclerosis. Hyperpigmentation and alopecia are found more commonly on ECDS but can be present in PRS. Histological examination of both shows evidence of dermal fibrosis in ECDS that is rarely present in PRS. Palpation of lesions has also been shown useful in differential diagnosis. In some cases of ECDS there is no sclerotic involvement but this is most commonly found in younger patients where the lesion has not been evolving for long and has a violaceous hue (Orozco-Covarrubias, Guzmán-Meza et al. 2002). Histopathological features are adnexal atrophy, mononuclear cell infiltrates and sclerosis. PRS and ECDS has similar ANA positive results (25% and 28.4%). In addition, the level of CNS and ocular involvement was reported to have similar levels.

In a global study of 205 people with PRS by Stone 2003, it was suggested that PRS and ECDS could be the same disorder and are very commonly found alongside one another. This online survey was the largest study of facial sclerosis but is based on the response of the patient directly. 21% of the participants reported the dual diagnosis of PRS and ECDS by their doctor. The study showed participants 3 pictures. Two pictures of cases of hemifacial atrophy of the lower face and a picture of ECDS. 31% of participants reported having both the “line” of ECDS and the lower face hemiatrophy suggesting a possible relationship or overlap of the diseases.

This study also asked participants where on the face they were affected and where else on the body. All had PRS of the face, 63% had it on the forehead and 51% of these participants' PRS had a linear pattern on the forehead. In 10% of participants, the linear lesion touched the corner of the eye, in 20% the lesion extended to the mid-eyebrow and 19% of patients' lesions extended to within one centimetre of the mid-line of the face. The cheek (75%), chin (43%), lips (55%), teeth and gums (50%) and tongue (25%) were also frequently affected. Interestingly, 19% of participants reported involvement of arms, legs and/or trunk alongside the PRS. This indicates not only a possible relationship between PRS and ECDS subtypes but also a connection between PRS and LLS. It must be stressed however that this study had no control group and there is likely some responder bias (Stone 2003).

In a study of 54 patients with PRS and ECDS it was hypothesized that the two disorders share a similar pathogenesis and both should be included in the banner head of linear morphea, but that in fact they are distinct diseases. In this study, 53.6% of the patients who were diagnosed with PRS were also shown to have ECDS lesions. More studies reported the incidence of both diseases appearing alongside were 42% and 41.7% respectively. Conversely the patients with ECDS, 36.6% of them were shown to have PRS (Falanga, Medsger et al. 1986, Sommer, Gambichler et al. 2006).

Many CNS related illnesses, mainly headaches and seizures have been found associated with ECDS and PRS in numerous studies. In (Stone 2005) 11% of PRS patients reported a diagnosis of epilepsy. In another study of PRS and its relationship to the CNS (Blaszczyk, Królicki et al. 2003), 21% of PRS patients (with and without co-diagnosis of ECDS) reported seizures. Another study reported 13% of PRS, ECDS and both had seizures (Tollefson and Witman 2007). Zulian found 8% of ECDS and PRS patients had seizures (Zulian, Vallongo et al. 2006).

Computed Topography (CT), Magnetic Resonance imaging (MRI) and EEGs have all been used in the search for brain abnormalities in PRS and ECDS patients. Intracranial abnormalities have been described in patients with seizures (Blaszczyk, Królicki et al. 2003, Zulian, Athreya et al. 2005).

It has been suggested in the case of both ECDS and PRS that they both follow the pattern of the trigeminal nerve in terms of skin presentation. As shown earlier by Resende *et al*. 1991 the SNS could possibly be involved whether it is hyper- or hypodysfunction of the SNS along the trigeminal nerve. This would suggest they are the same spectrum of one disease affecting certain trigeminal nerve branches with varying severities and therefore variation in cutaneous involvement. This might also account for the neurological and ocular complications in both, the high incidence of them appearing together and the difficulty in differential diagnosis.

8. Misdiagnosis

One of the most daunting aspects of LLS and morphea for patients is the high rate of misdiagnosis and the typical delay of several years before the correct diagnosis is reached. In a study based on children with morphea at Great Ormond Street Hospital consisting of 50 patients, the delay to diagnosis was analysed. The average age of the patients in the study was 5.2 years (0.1-14.4). 86% of the patients had LLS and 8% had LLS in combination with circumscribed morphea. From families of the patients it was determined that the average delay before doctor consultation after the first symptoms were noticed was 1.2 months (0.2-48.7). Patient families first suspected bruise, burn, atopic dermatitis, fungal infection, vitiligo, insect bite or other unknown causes (Weibel, Laguda et al. 2011).

The average delay of referral by the family doctor to a specialist i.e. dermatologist, was 7.5 months (1-70.9 months). For 64% of patients it was at this point that a correct diagnosis was made, the rest had no diagnosis or a misdiagnosis (a common misdiagnosis being port-wine stain) (Nijhawan, Bard et al. 2011). Overall it was reported that the delay to diagnosis after disease progression was averaging 11.1 (1.8-70) months (Weibel, Laguda et al. 2011). Others studies report delay in diagnosis at 18 months with 20% of patients waiting over 24 months (Zulian, Athreya et al. 2005).

Common differential diagnoses of morphea is erythema chronicum migrans, granuloma anulare, cutaneous mastocytosis, drug reaction, lichen sclerosus and others. The hyperpigmentation in morphea can suggest differential diagnoses such as Café-au-lait spots, post-inflammatory hyperpigmentation and lichen planus actinicus. Atrophy can also be compatible with scarring, lichen sclerosus and acrodermatitis chronica atrophicans. Sclerosing skin occurs with pretibial myxedema and necrobiosis lipoidica. ECDS is often

confused with panniculitis, focal dermal hypoplasia, lupus erythematosus profundus, steroid atrophy and panniculitis (Kreuter, Krieg et al. 2009).

9. Treatment

There are three main treatment types recommended in morphea. Topical drugs, immunosuppressive systemic compounds and phototherapy. Treatment is much more successful when the disease is diagnosed early. However, as is evident from literature and clinical experience, early diagnosis can be rather difficult. Many variables must be considered when treating the disease. These variables fall into two main categories. First is disease activity.

- Has a new lesion appeared in the last 3 months?
- Has a pre-existing lesion increased in size?
- Is there moderate to severe erythema or erythematous borders?
- Has there been use of imagery to document disease activity or progression?
- Has induration of the lesions increased?
- Has alopecia gotten worse?
- Is the skin biopsy showing an active disease?
- Is there an increase in creatine kinase (CK)?

Second is clinical damage

- Is there atrophy of the dermis?
- Is there atrophy in subcutaneous tissue?
- Is there dyspigmentation of the tissue?
- Is there increased thickness in the centre of the lesion?

(Careta and Romiti 2015)

Many treatments for LS and LLS have been used with varying and often disappointing results. Immunosuppressors methotrexate (MTX) and cyclosporin A (CyS) are commonly used with the aim to dampen the overactive immune system in patients and to reduce fibrosis of the skin. It is a first-line treatment in up to 50% of cases of LLS. In a study of roughly 460 LLS patients MTX was used as treatment in 44% of patients and CyS in only 2% (Zulian, Athreya et al. 2005).

A study of 34 patients with active LS lesions (Weibel, Sampaio et al. 2006) treated with a combination of corticosteroids (CS) e.g. prednisone and MTX showed that in 94% of patients the disease stopped progressing. Over time, the disease improved in all patients. Treatment was stopped in 47% of patients when the lesions were considered inactive. However, 44% of these patients relapsed, resulting in renewed treatment. Side effects were described as minimal and transient. At completion of the study, 71% of the patients had inactive lesions (Weibel, Sampaio et al. 2006, Kreuter, Krieg et al. 2009).

According to a survey from the Childhood Arthritis and Rheumatology Alliance (CARRA) of 70 clinical centres, focusing of pediatric rheumatology, a combination therapy of CS and MTX is the favoured treatment in North America. Of the 650 clinical cases that had been treated by doctors who participated in this study nearly all were treated with combination therapy. A major issue with this treatment is that no large-scale clinical trial has been performed to investigate the correct dosage and treatment time. Also there is no information about efficacy and response to treatment between different subtypes of morphea (Li, Feldman et al. 2010).

In a 2012 study of 36 patients to determine to best dosage of MTX and CS therapy it was found that an initial high dose of subcutaneous treatment of MTX (1 mg/kg/week max 25 mg/week) for two years was optimal. Subsequently the treatment method was altered to the

same MTX dosage but to an oral vector for 6 months followed by dose reduction over the next 6 months provided the disease was inactive. This was combined with oral prednisone daily (2 mg/kg/day tapered to 0.25 mg/kg/day). What they found was that the lesions stayed in remission for the entirety of the treatment in all patients (Torok and Arkachaisri 2012). Despite the promising results, this study was not a randomized, placebo-controlled study unlike the study by Zulian, *et al.* 2011. This double-blind study randomized 70 morphea patients 2:1 for oral MTX treatment versus a placebo. Treatment began with MTX (15 mg/m^2/week max 20 mg) for 12 months and a short course of CS (1 mg/kg/day) for 12 weeks. It was found that any adverse side effects were tolerable and lesion size decreased significantly though some patients (n=3) developed new lesions (Zulian, Martini et al. 2011).

Topical, oral and intravenous steroids are also given to patients in an estimated 42% of cases of LLS to treat inflammation. In Zulian *et al.* 2005 topical steroids were used in 13% of patients and oral steroids in 28%. A common treatment of ECDS in the clinic is triamcinolone acetonide mixed with lidocaine (10-40 mg diluted 1:2 to 1:4) injected intralesionally however there has been no studies determining its effectiveness. Systemic corticosteroids has shown some success when used alone or in combination with other therapies (Zulian, Athreya et al. 2005, Weibel, Sampaio et al. 2006, Kreuter, Krieg et al. 2009). One study on the use of steroids in a monotherapeutic approach in ECDS patients for 17 months showed some improvement however, up to a third of these patients relapsed after therapy ended (Joly, Bamberger et al. 1994).

D-penicillamine, a chelator, is also used in up to 30% of patients. However, there is no proof that it is effective at any dose and there is a high level of side effects reported (50%). These side effects can be quite severe ranging from rashes, diarrhoea, nausea and leukopenia (low

white blood cell counts). The benefits of this treatment and it effectiveness are therefore debated with much scepticism about whether or not it should be prescribed (Zulian, Athreya et al. 2005, Kreuter, Krieg et al. 2009).

Other, less common, therapies used are psoralen and ultraviolet A (PUVA), vitamin D (Calcipotriol) and NSAIDs. Low-dose PUVA phototherapy in combination with topical treatment of calcitriol has been shown to have successes when the disease is localized to the skin. There was a noticeable reduction in fibrosis and decrease in hyperpigmentation that maintained after follow-up appointment a year later. Mild skin irritation was reported from the phototherapy, however there was no other side effects (Gruss, Von Kobyletzki et al. 2001, Kreuter, Gambichler et al. 2001). This treatment has been found to be successful in softening the affected skin in ECDS in a limited number of patients (Gambichler, Kreuter et al. 2003). Treatment with the antifungal compound griseofulvin, colchicine and immunosuppressive drugs (Orozco-Covarrubias, Guzmán-Meza et al. 2002). Treatments do not reverse the effects of the disease, however it can sometimes be helpful in stopping the disease spread.

Mycophenolate mofetil (MMF) has also been shown to be beneficial as an alternative treatment to MTX and CS, as some patients can be resistant or intolerant to the treatment. It has been shown to soften the affected skin (Marsol 2013). In a study of 10 patients, a dose of 600-1200 mg/m^2 given twice daily. All patients showed an improvement in lesion size and sclerosis. There is, however no in-depth analysis of its efficacy but future studies can determine its usefulness (Martini, Ramanan et al. 2009, Mertens, Marsman et al. 2016).

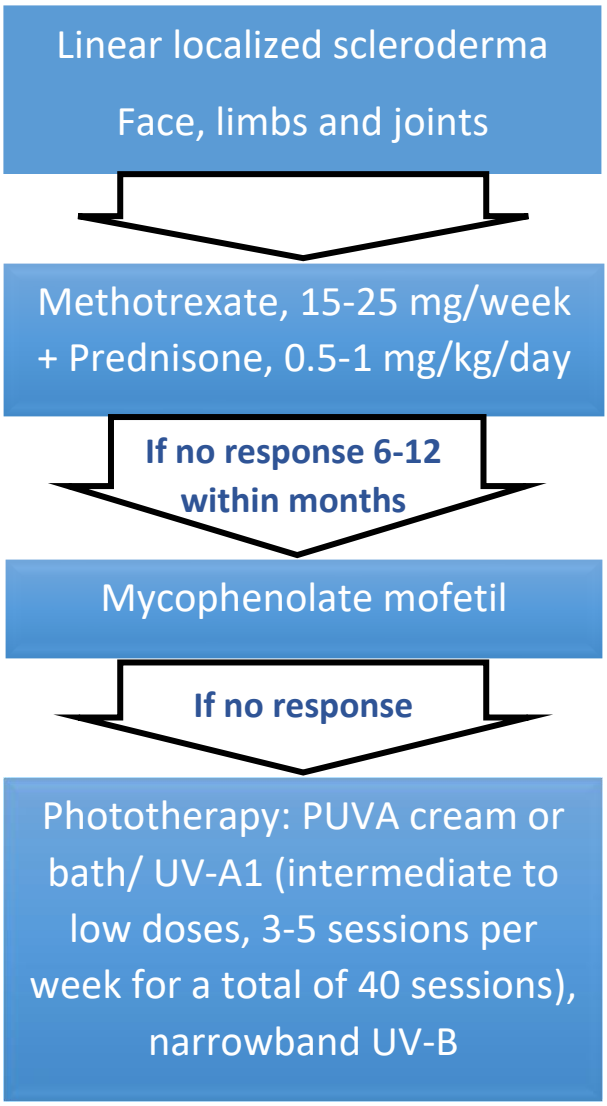


Figure 9. Adapted from Marsol. A proposed sequence of treatment of patients with LLS (Marsol 2013)

Treatment of LLS as suggested by Marsol above is quite different to other forms on morphea. In the case of generalized morphea, phototherapy is the preferred first treatment followed by prednisone and methotrexate and finally MMF (Figure 8). Plaque morphea is a more limited and superficial subtype of morphea and thus topical treatments are nearly exclusively recommended in these cases (Marsol 2013, Careta and Romiti 2015).

Surgery is often sought as a treatment, as stable scarred lesions of sufficiently small size can be removed by excision. This treatment is mostly reserved for adults as it is of the utmost importance that the lesions are completely inactive. A period of several years stability without inflammatory events suggests clinical inactivity. In some cases, waiting for disease inactivity is not possible, for example in cases where differences have arisen in arm or leg length due to lesions inhibiting growth. Epiphysiodesis must be performed during childhood regardless of disease stage or activity. If this is the case, the patient growth must be monitored hereafter in combination with a systemic immunosuppressive treatment (Kreuter, Krieg et al. 2016). Cosmetic surgeries have been used in ECDS and PRS treatment. Autologous fat grafting, insertions of fillers and removal of affected skin are all surgical options but must only be performed in patients with inactive lesions (Figure 9) (Karaaltin, Akpinar et al. 2012).

There is as of yet no agreed nor perfect treatment for morphea and LLS. It appears to be case dependent. However, the most advantageous tool is an early diagnosis. There are numerous treatments the most successful ones being PUVA treatment sessions or MTX with CyS treatment.

Figure 10. Surgical reconstruction in a patient with PRS A) PRS affecting the left cheek B) post-surgical reconstruction (Raposo-do-Amaral, Denadai et al. 2001)

10. Differentiation from Systemic Scleroderma (SSc)

The term “scleroderma” is the hardening and thickening of organs due to deposition of collagen. In morphea the sclerosis is limited to skin and subcutaneous tissue sometimes with some extracutaneous involvement. SSc is the most severe connective tissue disorder. Once it presents it becomes debilitating and often fatal with the poorest prognoses for patients with early onset. It is a multisystem disease characterized by immune dysregulation, visceral and cutaneous fibrosis and vasculopathy. SSc has two categories based on pattern of skin involvement, limited cutaneous sclerosis and diffuse cutaneous sclerosis. 70% of patients have pulmonary involvement, which is the highest cause of mortality (Abraham, Krieg et al. 2009, Elhai, Meune et al. 2011, Ramos, Silver et al. 2015). Other organs can become involved such as heart, gastrointestinal tract and kidneys.

Aside from visceral organ involvement, SSc is differentiated from morphea by symptoms like Raynauds phenomenon, capillaroscopic abnormalities and sclerodactyly. Both diseases have symptoms such as malaise, myalgia and fatigue and sometimes positive autoantibody serologies (Fett and Werth 2011, Amaral, Marques Neto et al. 2012). No specific serological markers for LS are available as of yet. Interleukins 2, 4, 6 and 8, and surface molecules CD23 and CD30 have previously been reported in some cases of LS but no substantial relationship has been shown (Kreuter, Krieg et al. 2009). There are many genes that have been found to be associated with SSc. Most of the genes that have been implicated in SSc are immune genes that modulate immune reactions (Ramos, Silver et al. 2015). As of yet, no studies have attempted to find a genetic link between SSc, morphea and LLS.

Somatic Mutations and Mosaicism

1. Blaschko's lines

Blaschko's lines (BL) are the observed patterns of lesions in cutaneous mosaicism disorders. The patterns were first described in 1901 by German dermatologist Dr. Alfred Blaschko. He analysed 150 patients with epidermal and sebaceous nevi and made a composite diagram of the distribution on the body showing "S" shapes on the abdomen, "V" shapes on the back and arcs on the chest (Figure 11A) (Blaschko 1901, Molho-Pessach and Schaffer 2011). The BL model was then expanded upon by Dr Rudolf Happle in 2001 when he analysed and included the patterns on the head and neck (Figure 11B). 186 pictures of patients with lesions on the head and neck were compiled and again a composite diagram was made. He described "V" shapes on the neck, spirals on the scalp, perpendicular lines with intersection on the face and an hourglass pattern was found on the nasal ridge (Happle and Assim 2001, Molho-Pessach and Schaffer 2011).

These dermal patterns represent the coordinated proliferation, migration and apoptosis of ectodermal cell during normal embryological development. The patterns are always present in each individual and remain invisible in healthy skin but will be apparent in diseases that follow these lines (Biesecker and Spinner 2013).

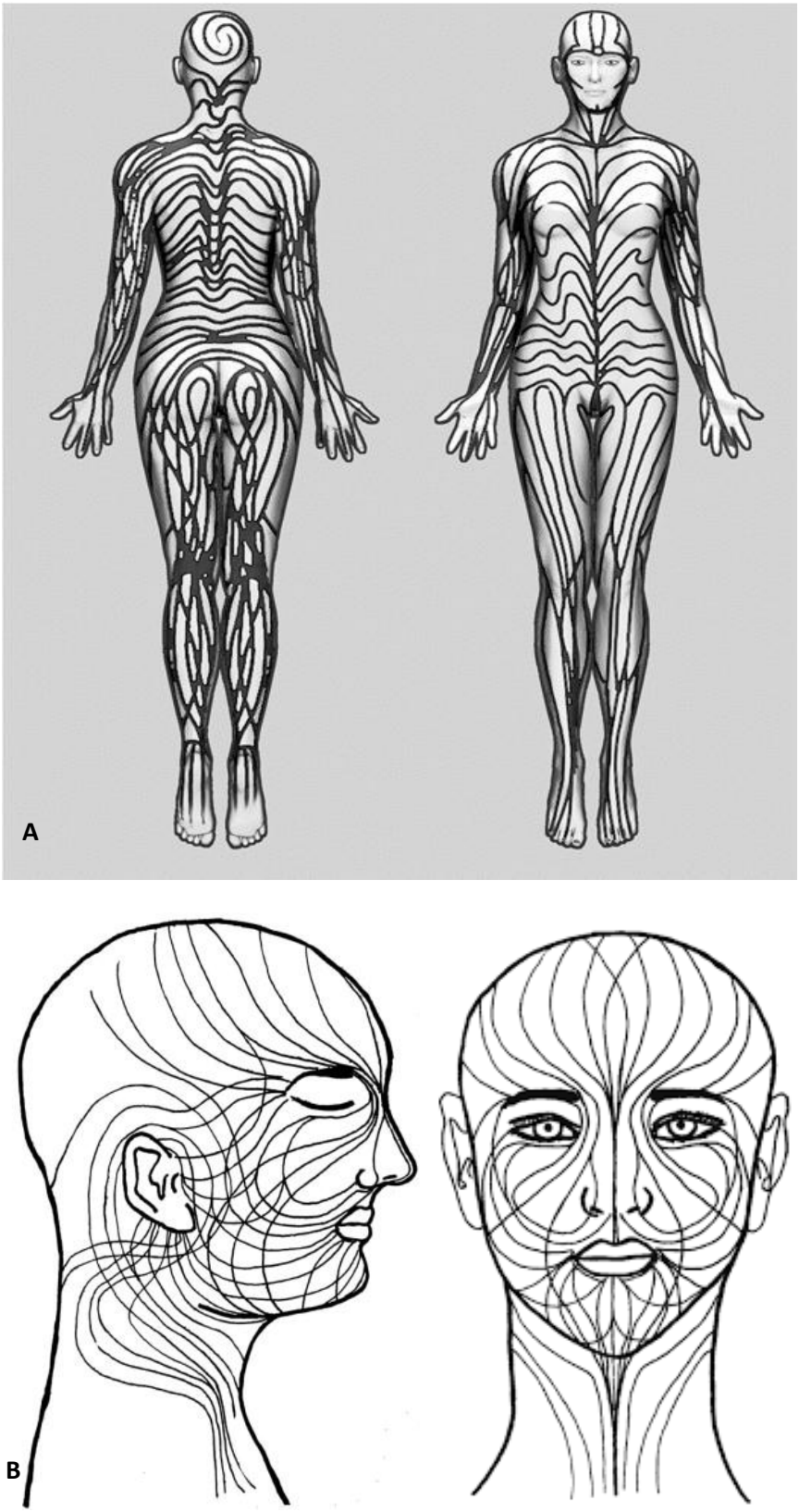


Figure 11. Blaschko's lines on A) the body, front and back (Blaschko 1901) and B) the head (Happle and Assim 2001)

2. Germ layers of the embryo

After fertilization a zygote forms. The zygote undergoes cell division and cleavage, making genetically identical cells in a hollow ball with the outside consisting of a single layer of cells. This is called a blastula. The blastula then undergoes gastrulation to create a multi-layered structure called a gastrula, which contains the germ layers (Figure 12). In humans three germ layers form endoderm, mesoderm and ectoderm. Christian Pander determined in the 1820s that from these three layers all the organs of the embryo will arise.

Ideally, during embryogenesis all of the cells remain genetically identical but often mistakes in mitosis occur. Though there are some fail safes to correct mistakes in DNA duplication some mutations remain. If a mutation arises in one cell, all the descendants of this cell will contain the mutation. At the early stages of embryogenesis there are fewer cells so the earlier the mutation arises the higher the number of cells that will contain the mutation. This alteration of the genotype results in the presence of two genotypes in one organism. This is called mosaicism.

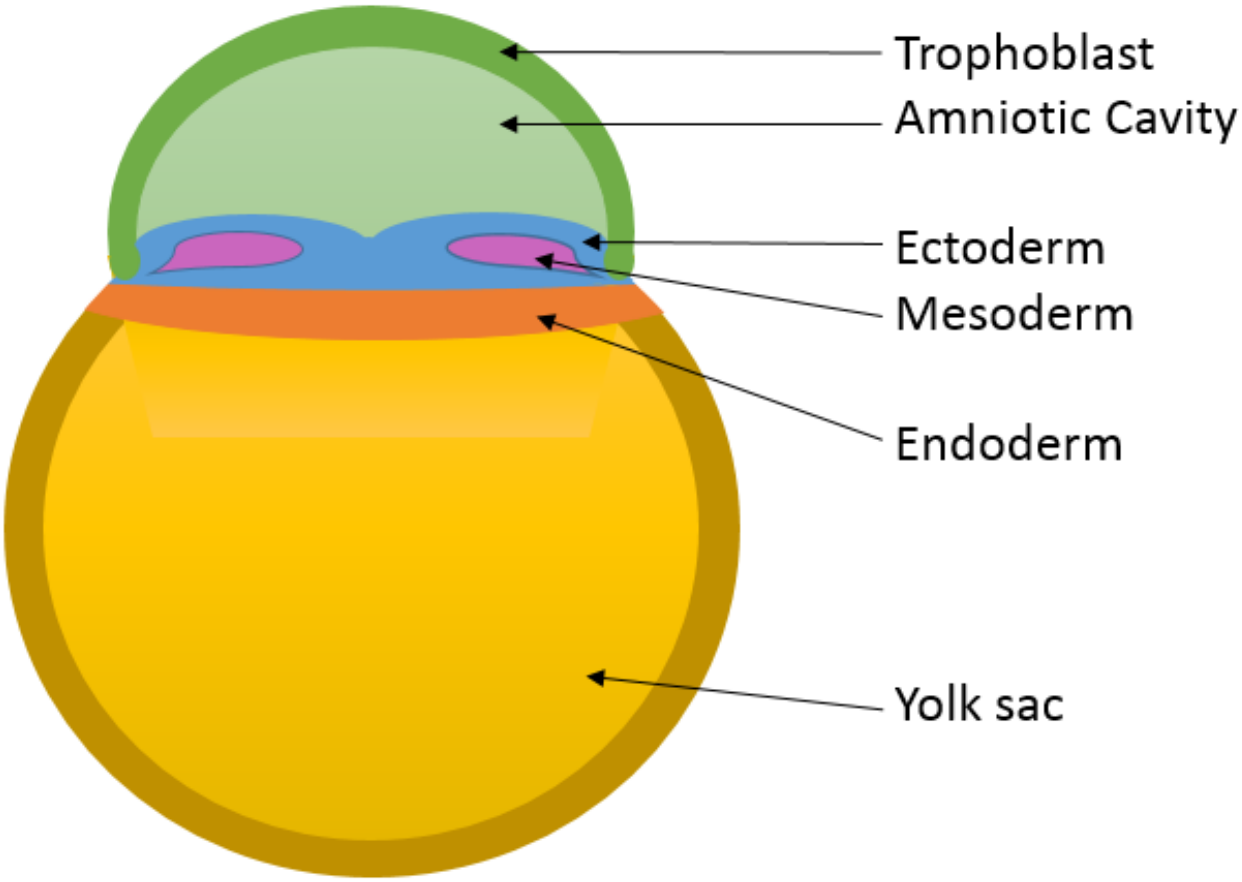


Figure 12. Embryo at 16 days (gastrula) showing origin of three germ layers.

3. Mosaicism

Mosaicism is named after the art of using different pieces of coloured glass and tiles to make patterns or images. It is the presence of two genotypes that originated from a single homogenous zygote (Paller 2004). It is differentiated from chimerism, which also has two different genotypes but arises after the fusion of two zygotes leading to two separate cell lines in one individual. Mosaicism is quite rare but the most widely known examples are in phenotypes such as heterochromia iridum (the difference in colour of eyes, skin, or hair), the variations in colours of dog coats and the most infamous mosaicism of all, cancer.

Cutaneous mosaicism is also found in dermatological diseases in humans that follow Blaschko's lines. Mosaicism can be germline (gonadal), somatic or a combination of the two called gonosomal. Germline mosaicism is when the gametes of a person carry the mutation, while the rest of the cells do not. This will affect the children of this individual. Somatic mosaicism is when the mutation is found in the some i.e. cells that are not germ cells. For this reason, these mutations are not inheritable and will arise spontaneously (Figure 13) (Biesecker and Spinner 2013).

There are two types of cutaneous mosaicism. The most common form is type I. This is when the mosaic skin is linear and follows Blaschko's lines but uninvolved, surrounding skin is healthy. This type of disorder was first demonstrated in patients with epidermal nevi. Keratin 19 mutations were found in affected Blaschkoid skin, whilst there was no such mutation in the normal skin. Histological analysis of affected skin showed stratum corneal thickening and epidermolysis not present in healthy skin.

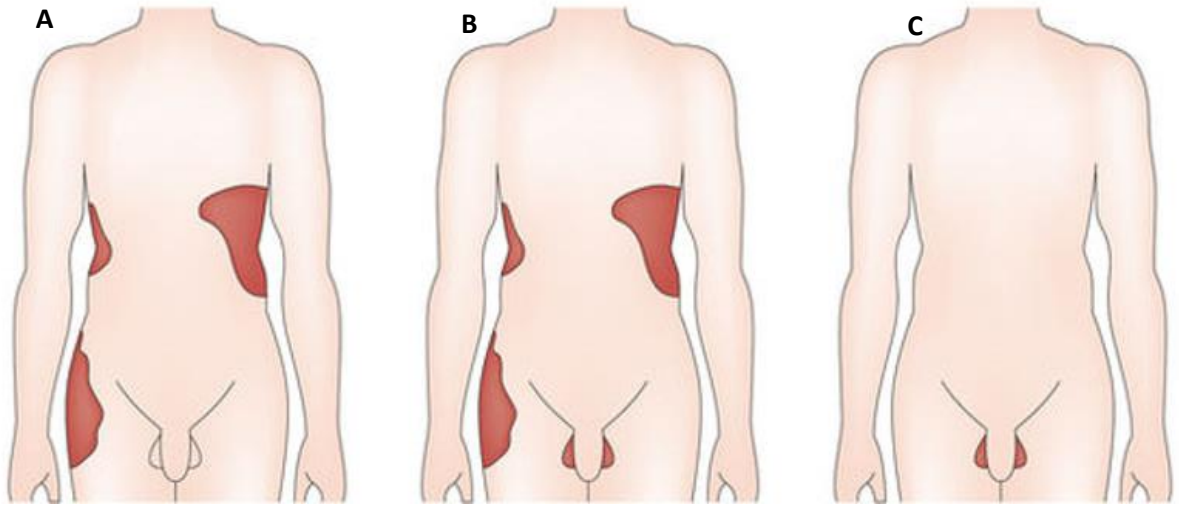


Figure 13. Types of mosaicism A) Somatic, arising spontaneously in a tissue, not heritable
B) Gonosomal, affecting some tissues including gametes resulting in them possibly being passed to offspring
C) Gonadal, present in gamete or sex organs only (Biesecker and Spinner 2013)

Type II cutaneous mosaic has been described. This is when the lesions of a disease follow Blaschko's lines and its effects are severe but there is diffuse skin involvement to a lesser extent. Poblete-Gutiérrez *et al*. 2004 have described this where they studied a patient with Hailey-Hailey disease, an autosomal dominant disorder. On analysis of severely affected skin versus surrounding less affected skin, they found a mutation in the *ATP2C1* gene was present in both. In the severely affected skin there was a loss of heterozygosity event. This means the heterozygous mutation was found in affected skin of the patient but where there was a mosaic loss of heterozygosity, the effect of the mutation was doubled (Paller 2004, Poblete-Gutiérrez, Wiederholt et al. 2004).

4. Mosaicism in Mendelian Disorders

Mendelian disorders are the most widely known and understood pattern of inheritance. This is where an affected parent or both parents pass a disease-causing gene onto their children via any of three possible ways; autosomal dominant, autosomal recessive and X-linked. Most disorders are inherited in this way. The mutations that cause these diseases can more rarely arise *de novo* in a somatic pattern. Several Mendelian disorders have been described in a mosaic form.

4.1 Apert syndrome

Apert syndrome (AS) is a severe disease characterized by craniosynostosis, mid-face hypoplasia and syndactyly. This means the patients have fused prematurely fused cranial plates, an underdeveloped mid-face and fused fingers, resulting in severe disfigurement. The premature fusion of skull bones is also detrimental to intellectual development. A diagnosis is generally achieved in early life. Another feature of AS is severe acne, which generally develops during puberty (Henderson, Knaggs et al. 1995). Genetic analyses of patients with AS have unearthed various *FGFR2* mutations in patients. *FGFR2* (fibroblast growth factor receptor 2) is an important gene in numerous processes including cell growth, wound healing and embryonic development. Mutations in *FGFR2* have been associated with many severe congenital cytoskeletal disorders (Kiritsi, Lorente et al. 2015).

In one case of severe acne in a 12-year-old boy appearing on previously hypopigmented lesions following Blaschko's lines. DNA was isolated from blood and affected skin. *FGFR2* was sequenced using Sanger sequencing. A mutation of Ser252Trp of *FGFR2* was found in the skin

but not in the blood. This mutation has been previously described in patients with AS. This indicates a post zygotic mutation occurred and explains the lack of skeletal involvement in the patient. The severe acne was treated with oral isotretinoin, a chosen treatment for the acne of AS. This left the patient with scarring, replacing the comedones and inflammatory papules (Kiritsi, Lorente et al. 2015).

4.2 Darier's Disease

Darier's Disease (DS) also known as Darier-White or keratosis follicularis is an autosomal dominant disease caused by mutations in *ATP2A2*. It is characterized by a scaly or bumpy rash found symmetrically throughout the body. It is caused by a malfunction of keratinization disrupting cell adhesion causing acanthosis. It can range from mild to severe with the milder forms often going undiagnosed. Symptoms generally appear during adolescence and will advance to complete penetrance as an adult. *ATP2A2* codes for sacro-endoplasmic reticulum Ca^{2+}-ATPase type 2 pump that is found in keratinocytes and muscle cells. An estimated 140 mutations of *ATP2A2* have been associated with DS.

Mosaicism in DS has been reported several times and is called segmental DS. Both Type I and II mosaic manifestations of DS have been described. Few studies have investigated segmental DS of yet. One such study by Harboe *et al*. 2011 found a mutation in *ATP2A2* in a male patient with type I segmental DS with no familial involvement (Harboe, Willems et al. 2011). In 3 other similar studies involving a total of 4 patients, 3 patients were found to have mutations in *ATP2A2* (Sakuntabhai, Dhitavat et al. 2000, Wada, Shirakata et al. 2003, Huh, Fujiwara et al. 2007). There is some debate as to whether segmental DS is the same as acantholytic dyskeratotic epidermal nevi (ADEN). This is due to the presence of acantholytic lesions and

the observed worsening of symptoms after exposure to sunlight that is present in both diseases. Some studies have found mutations of *ATP2A2* in ADEN (Sakuntabhai, Dhitavat et al. 2000). However some studies maintain these should be differentiated (Wada, Shirakata et al. 2003, Huh, Fujiwara et al. 2007).

5. Mosaicism in LLS

In a retrospective study of 65 paediatric patients with LLS were chosen for a study based on the availability of clinical photographs (Weibel and Harper 2008). Using computer aided analysis, common patterns of the lesions were found. 31 patients had LLS on the trunk and limbs and 38 had either ECDS or PRS. The investigators found that LLS follows Blaschko's lines on both the body and the head. This study suggested that LLS is likely caused by a somatic mutation that causes a genetic mosaicism. If LLS is caused by a genetic mosaicism along BL, this could also explain why there are often more than one lesion and the incidence of co-diagnosis of LLS with PRS. The mosaic cells could therefore be susceptible to LLS and then triggered by some factor that is still unknown (Weibel and Harper 2008, Careta and Romiti 2015).

Unlike most common types of mosaicism, morphea does not occur in ectodermal structures (section 2.3). All subtypes of morphea, including LLS, cause sclerosis of mesodermal-derived tissues (Paller 2004, Fett and Werth 2011). Cases like this have been described before in atrophoderma of Moulin or Goltz syndrome, where the affected areas follow Blaschko's lines, but the disease affects the dermis (Weibel and Harper 2008). This explains why there is no

pattern of inheritance of LLS when family history of patients is investigated (Christen-Zaech, Hakim et al. 2008, Fett and Werth 2011).

Genetic investigation of somatic mutations

1. Background

Patients affected with any given disease do not usually demonstrate the exact same pattern of signs and symptoms. The somewhat unpredictable variability is the main difficulty for diagnosis of many conditions. Variability can be due to many factors that are specific to each individual. Such factors include the environment, modifier genes, the epigenome and the synergistic effect of other mutations in the patient. This results in variation in penetrance between patients. High penetrance indicates the disease has a strong effect on the phenotype and is more associated with familial Mendelian disorders, whereas low penetrance have a weaker disease phenotype and are associated with diseases with more complex traits (Voelkerding, Dames et al. 2009).

Genomics comprises of the study of the genome of any organism to understand its structure, function and sequence. This focuses on the genetic code as a whole rather than focusing on a single gene at a time. The most important advancement in the field of genomics was the invention of DNA sequencing. Through large-scale analysis of DNA it is not only possible to diagnose diseases but it bridges that gap between genotype and phenotype.

2. The Human genome

The human genome has roughly 3 x 10^9 nucleotide bases. This includes both coding and non-coding regions. The first genome was published in 2004 by the International Human Genome Consortium. It cost an estimated 300 million USD but was one of the most significant scientific leaps forward. This also became the first reference genome. 3 years later Craig Venter and his colleagues published the second completed genome. Comparison of the two revealed 3.2 million SNVs (Metzker 2010). This is more than 0.01% variation between two genomes. Since then, new references genomes have been published and with the advent of next generation sequencing technologies (NGS) they have become more comprehensive.

The reference genome Genome Reference Consortium Human Build 37 (GRCh37) (Feb. 2009) from University of California Santa Cruz (UCSC) also known as hg19 was used as a reference for all analysis and assembly in our study (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/). According to Ensembl (http://grch37.ensembl.org/Homo_sapiens/Info/Annotation#assembly) this reference release includes 27478 contigs (overlapping sequence reads), 9 haplotypic regions of MHC classification (genes inherited as a group) and the chromosome length is 3.1GB. It therefore currently includes 20,805 protein-coding genes, non-coding genes and known pseudogenes.

Large-scale data collections of sequence variants such as the 1000 Genomes Project (TG) (http://www.internationalgenome.org/) and Exome Sequencing Project (ESP) (http://evs.gs.washington.edu/EVS/) have unearthed the broad spectrum of human genetic variation. These projects have analysed genomes and exomes from all over the world and determined the expanses of mutations that are commonly found and, more importantly, uncommonly found. The projects have also gone one-step further with subgroupings of

datasets in geographic specific areas such as Europeans, American, Asian and Africans. This allows elimination of common sequence variants in the patient according to their genetic heritage and to focus on the rare, or as of yet, unclassified mutations. This proves invaluable tool for bioinformaticians when analysing patient data for disease characterization at clinical level.

3. DNA sequencing

The earliest form DNA sequencing was developed at Cornell University in the early 1970's. The basis of this technology involved the use of DNA polymerase and nucleotide labelling which are still used in modern-day sequencing technologies (França, Carrilho et al. 2002). Improvements to the technique resulted in faster and more comprehensive methods known as NGS technologies (Mardis 2008). These technologies come from Illumina, Roche 454 and Ion torrent amongst others and they can sequence DNA with several techniques. The most utilized techniques are discussed here.

Whole exome sequencing (WES)

WES focuses solely on the protein-coding region of the genome. WES has proven very useful at the clinical level by allowing for swift implication of mutations in diseases (Bolze, Byun et al. 2010, Gilissen, Arts et al. 2010). It can be useful to clinicians when dealing with symptoms with no apparent or obvious connection to a disease or a disease with an abnormal clinical presentation. WES has proved an invaluable tool for rare diseases where a more personalized

medicine approach must be adopted in the treatment of a patient (Dyment, Smith et al. 2015). WES methods are described further in the Materials and Methods section.

Whole genome sequencing (WGS)

WGS is a technique where the whole DNA sequence of an organism is determined. It is the most comprehensive analysis of the DNA sequence. Much like WES, it has been useful in identifying disease-causing mutations in several genetic disorders and is an invaluable method in genomics (Saunders, Miller et al. 2012). This method is similar to WES as it uses fragmentation and library amplification thought PCR. However, it differs in that introns and exons are not separated resulting in a more extensive analysis (Mardis 2008).

Sanger sequencing (SS)

SS was developed by and named for Dr. Sanger and his colleagues in 1977 (Sanger and Coulson 1975, Sanger, Nicklen et al. 1977). This method was heavily utilized in sequencing the first human genome. It remains one of the most highly used form of DNA sequencing. It works by producing a complementary strand of DNA using DNA polymerases and deoxynucleoside triphosphates (dNTP) monitoring the incorporation of modified nucleoside, dideoxynucleoside triphosphate (ddNTP) into the growing chain (França, Carrilho et al. 2002). SS laid the groundwork for future sequencing techniques including WES and WGS.

4. Choice of DNA sequencing technique

In this study, WES was the preferred technique for several reasons. For one, the exome is estimated to be 1% the size of the genome at roughly 30 megabases. Due to the smaller size of the exome, a higher average coverage can be achieved with less cost and less raw sequence. Due to the targeted nature of WES takes 3GB of raw sequence data to achieve 75x coverage whereas with WGS it takes 90GB of raw sequence data to achieve 30x coverage.

It has been estimated that 85% of disease-causing mutations are found in the exome (Bao, Huang et al. 2014). Currently less than 10% of the whole genome has been characterized, meaning there is limited clinical gain to sequence the full genome. In addition to the lower cost of WES when compared to WGS, there is a lower computational resources required in WES analysis and thus a quicker data turn around (Majewski, Schwartzentruber et al. 2011).

SS was not chosen for this project because in this technique one can only sequence a region or gene which is known to be of interest. In this study, the area of interest is unknown as LLS has not been studied from a genetic standpoint before. It is for this reason that SS was concluded to be a non-viable candidate for analysing the LLS patients.

However, SS still has an important role as it is used as a technique to confirm mutations found using WES and WGS. Confirming a mutation with two forms of sequencing techniques strengthens the argument for the presence of the mutation.

There is a disadvantage to the targeted sequencing process of WES. The most important step is the separation of the exons from the introns. This is achieved using a “targeted” approach where oligonucleotides (short RNA molecules) are designed that attach solely to exons. This approach allows the researcher to design their own targets according to the specific genes of

interest or genes that are known to be involved with diseases. This is especially helpful for those who wish to screen for cancers. However, target capture is often unequal in that there is inefficient capture rates across the exome leading to preferential coverage of some areas and low coverage in others. In addition, off-target hybridization can lead to capture of intronic regions and can lead to over 20% of reads to be intronic reducing efficiency (Voelkerding, Dames et al. 2009).

Sequencing was performed on the Illumina HiSeq platforms. Illumina was chosen for two reasons. Firstly, it offers paired-end sequencing. This is a technique where reading of the DNA by the sequencer is performed in both directions. This is helpful in unearthing genomic rearrangements and repetitive sequences. Secondly, a single flow cell can produce up to 500GB of data per run. For this reason, it is possible to analyse several samples per run with high coverage.

Sequencing on the Illumina sequencer begins with attachment of the prepared library to the flow cell. During library preparation there is a step where adaptors for the flow cell are attached to the DNA. The flow cell is a glass slide with lanes coated with a lawn of 2 types of oligonucleotides. The DNA is hybridized by one of its adaptors to one of the surface oligonucleotides of the flow cell. A copy of the hybridized DNA is created by a polymerase and the original template is washed away. Next, bridge amplification takes place where the strand folds over and the adaptor binds to the second type of oligonucleotide. There is a copy made and this is then denatured resulting in two strands of copied DNA attached to the flow cell. This process is repeated and simultaneous reactions occur throughout the flow cell amplifying the number of copies of each DNA read.

Reverse strands are removed and forward strands are now sequenced. Fluorescently tagged nucleotides are added to the growing read. Each time a nucleotide is added the clusters are excited by a light source emitting a fluorescent signal. Wavelength and intensity of this signal is representative of the added base which is recorded by a camera. The forward strand then duplicated in a process similar to bridge amplification. The forward strand is removed and the reverse strands are now sequenced. All samples are separated according to the specific barcode given to each sample. This is now ready for bioinformatic analysis.

5. Formalin fixed paraffin embedded material for WES

Formalin is a favoured fixative in histopathology for examination and long-term preservation. It preserves the cells and their spatial relationship as well as carbohydrates and proteins amongst other large molecular structures. This allows detailed morphological analysis of the cell layers by a histopathologist, often an important step in diagnosis (Wong, Li et al. 2014). It is a preferred fixative because formalin penetrates the tissue easily and cross-links with tissue in a covalent reaction with minimal shrinkage and distortion of the tissue (Thavarajah, Mudimbaimannar et al. 2012). Formalin has been used in museum collection for long-term storage of samples for more than 150 years but this method of storage can cause DNA damage in many ways. Fragmentation and cytosine deamination can have negative repercussions on sequencing and can cause unreproducible sequencing artefacts (Figure 14) (Wong, Li et al. 2014, Hykin, Bi et al. 2015).

Fragmentation can be bothersome for many types of analysis, which requires largely intact DNA, for example, comparative genomic hybridization, however, for WES the fragment size

of the input DNA averages 150bp. This level of fragmentation is not often reached in FFPE samples. Therefore, even fragmented FFPE samples can be used successfully for WES.

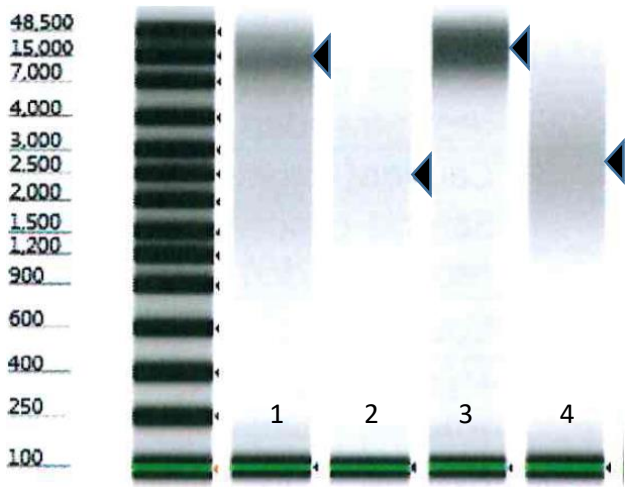


Figure 14. Genomic ScreenTape performed on ScreenTape System (Agilent Technologies, Santa Clara, CA, USA). Lanes 1 and 3 are DNA from SF sources; lanes 2 and 4 are DNA from FFPE sources. Arrows show fragment size. FFPE ~2,500bp and SF >15,000bp.

Sequencing artefacts are caused by the deamination of cytosine to uracil or methylated cytosine to thymine. This can result in cytosine nucleotides being reported as thymine (C→T) after sequencing. These errors can lead to false-positive results. One solution is to increase sequencing depth. It has been found that at an allelic frequency (AF) of 10-25% these C→T artefacts are less apparent as opposed to the disproportionate level found at AF 1-10% (Wong, Li et al. 2014). A more effective solution is the treatment of the FFPE DNA with uracil-DNA glycosylase (UDG) sometimes also called UNG. UDG is a repair enzyme, which locates uracil in DNA and hydrolysing the N-glycosidic bond between the base and deoxyribose. The base excision repair system will then repair this damage by replacing the base with the original cytosine nucleoside (Do and Dobrovic 2012).

Previous DNA analysis-based studies have included primarily or exclusively SF skin or FFPE skin. As part of this small clinical trial, skin samples received are either SF or FFPE depending

on when the patient was diagnosed and gave permission for inclusion. It was therefore important to make sure both preservation modes are of a high enough quality with negligible differences to be included in one study. SF samples can be kept in -80°C until DNA is extracted with some DNA degradation but no significant damage. The downstream effects of formalin on DNA must be avoided to prevent false positive mutations.

Materials and Methods

Patients and clinical samples

Skin samples, blood and saliva were received from patients as part of the LLS clinical trial (number: NCT02222038 https://clinicaltrials.gov/).

Inclusion criteria for patients are

- Male and female
- Between 5 and 30 years old
- Histologically confirmed and well phenotyped LLS by a clinician
- ECDS with or without therapy
- Lesion on any site of the body

Exclusion criteria

- Patients with signs of SSc (nailfold capillary, Raynaud phenomenon, sclerodactyly)
- Patients with any of the other 4 subtypes
- Patients with diagnosed gadolinium-induced scleroderma
- Patients with post-irradiation scleroderma
- Patients who have not given written consent

19 patients were included for analysis. Clinical data was available for 17 of these patients and data was recorded in Microsoft Excel 2013 (Microsoft Corp., Redmond, WA, USA) (Table 8-13). The clinicians took biopsies using local, topical anaesthetic cream (EMLA®) and a local injection of Lidocain 1% (1-2 mL). Patients with ECDS had an MRI to investigate the presence of intercranial abnormalities. If these patients were under 7 years or are experiencing anxiety, general anaesthetic was used and the biopsy was taken from these patients while anesthetized. Biopsy wounds were closed with self-dissolving or non-dissolving stiches.

A clinician collected 5mL EDTA blood from all patients. Local topical anaesthetic cream (EMLA®) was used in small or anxious patients to minimize pain and discomfort. Saliva was also collected from patients using Oragene DNA OG-500 (DNA Genotek Inc., Ontario, Canada) and both blood and saliva were sent to the laboratory at room temperature via courier.

Processing of biological samples

Skin samples received were either formalin fixed paraffin embedded (FFPE) (10% formalin: 3.7% formaldehyde in water and 1% methanol) or fresh punch biopsies taken by the clinician which were subsequently snap frozen (SF) (FFPE=9, SF=10) (Thavarajah, Mudimbaimannar et al. 2012). Biopsies taken after commencement of the study were between 4-5mm. They include dermis and epidermis of lesional skin only. Biopsies were halved with one-half to be FFPE embedded for histological analysis and the other half to be included in the study (~2mm). Fresh biopsies were placed in 1.0mL Cryotubes (Thermo Fisher Scientific, Waltham, MA, USA) and the whole tube was placed in liquid nitrogen to freeze them quickly and then stored long-term at -80°C.

Blood and affected skin were taken from 19 histologically confirmed LLS patients and WES was then performed pairwise. DNA was extracted from FFPE skin samples by taking 10x10µm slices using a microtome HM325 (Thermo Fischer Scientific, Hampton, New Hampshire, United States). DNA was then extracted using the GeneRead DNA FFPE kit (Qiagen, Hilden, Germany) with the protocol modified accordingly for increased input of slices. The suggested input is one slice at 10µm but due to the very small size of embedded skin, paraffin is instead trimmed away from embedded skin to facilitate deparaffinization. The kit involves a treatment step with uracil N-glycosylase (UDG), an enzyme shown to counteract cytosine deamination of DNA by formalin (Do and Dobrovic 2012).

Digestion of SF skin was performed by incubation in extraction buffer (1M Tris, 0.5M EDTA, Tween 20) and proteinase K (Roche, Basel, Switzerland) at 55°C overnight followed with a 10 min heat shock at 95°C to deactivate the proteinase K (Qiagen, Hilden, Germany). This digestion degrades skin complex protein structure and releases skin cells into the solution

increasing efficiency in DNA extraction. DNA was extracted from digested skin using QIAamp DNA blood mini kit (Qiagen, Hilden, Germany). DNA quality was checked using both a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

Peripheral blood mononuclear cells (PBMC) were extracted from EDTA blood using a density gradient. Blood was gently pipetted onto Ficoll-Paque PLUS (GE Healthcare, Little Chalfont, United Kingdom) in a 50mL Falcon tube (Fisher Scientific, Hampton, NH, USA) to form two layers (Figure 15A). Blood was centrifuged through Ficoll to separate blood according to density (2,500 rpm for 20min) (Figure 15B). The PBMC layer was extracted and cells are washed twice with phosphate buffered saline (PBS) pH 7.4 (Gibco by Life Technologies, Carlsberg, CA, USA) by centrifugation of the extracted PBMC and PBS mixture (1,500rpm for 5 min). Cells were resuspended in 200µL of PBS and DNA is then extracted from them using QIAamp DNA blood mini kit (Qiagen, Hilden, Germany).

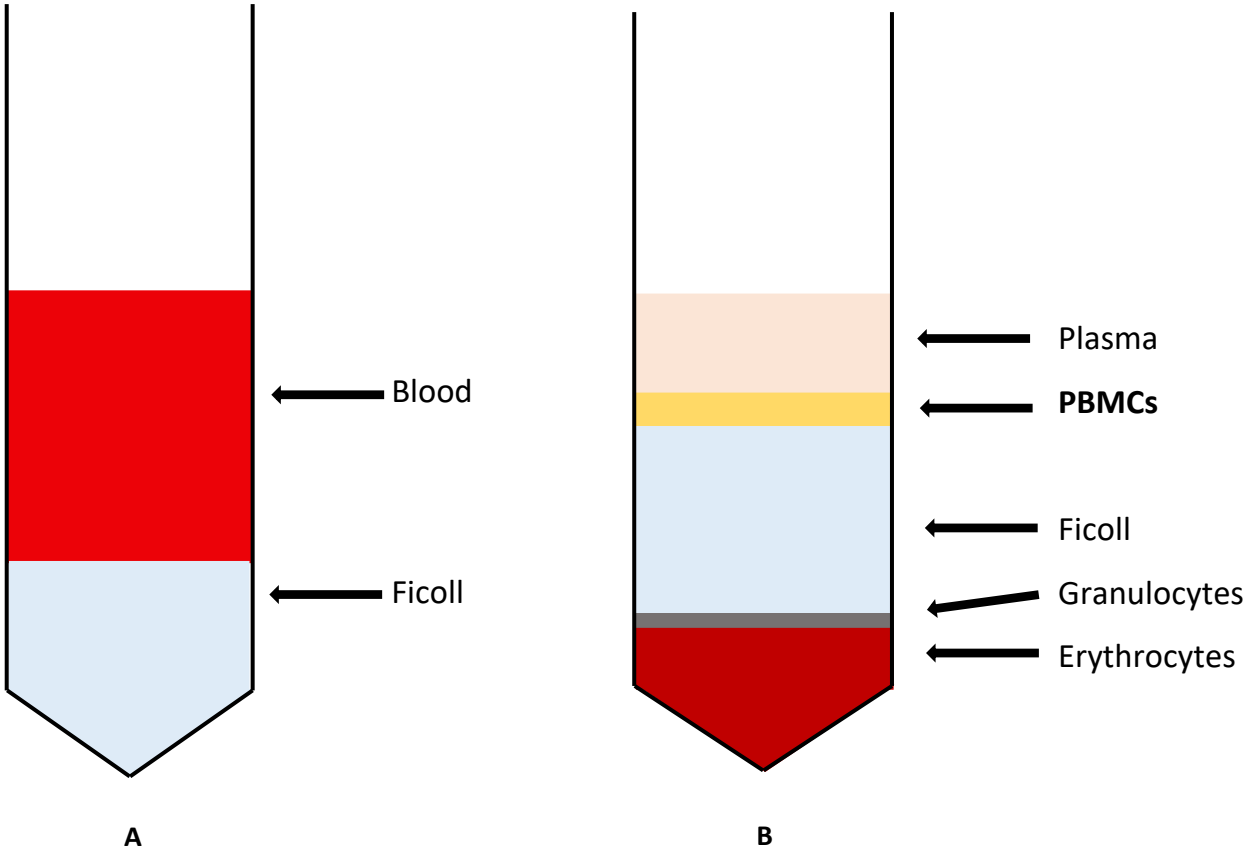


Figure 15. Ficoll extraction A) Layers prior to centrifugation. Blood pipetted on top of ficoll B) after centrifugation, cells are separated out according to density in a gradient

Whole Exome Sequencing

1. Library preparation

Input material was 200ng of genomic DNA. Quality and degradation of the genomic DNA was measured using the Genomic DNA ScreenTape System (200bp to > 60'000bp) on a TapeStation 2200 (both Agilent Technologies, Santa Clara, CA, USA). Samples were sonicated using a Covaris E220 to achieve a size of 150 - 200bp (Figure 16A) as per specifications in SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library Protocol (Version C0, December 2016)(Chen, Im et al. 2015). Size was confirmed using Agilent High Sensitivity D1000 ScreenTape System (Agilent Technologies, Santa Clara, CA, USA). End repair, adenylation and adaptor ligation was performed using SureSelectXT Reagent kits (Figure 16B-D) (Agilent Technologies, Santa Clara, CA, USA) and clean-ups between steps was performed using Agencourt AMPure XP magnetic beads (Beckman Coulter Inc., Brea, CA, USA).

End repair

The T4 DNA polymerase and Klenow enzyme are used to blunt the ends of the sonicated DNA. The 3' overhangs are removed and the 5' overhangs are filled in using the exonuclease activity in a 30 min incubation at 20°C.

Adenylation

'A' bases are added to the 3' end using the Exo(-) Klenow enzyme by incubation at 37°C for 30 min. This prepares the DNA for adaptor ligation.

Adaptor ligation

Single 'T' bases are present on the adaptors allowing them to be ligated to the DNA by DNA ligase for 15 minutes at 20°C. Adaptors allow the DNA to be attached to the flow cell, facilitate

primer attachment during PCR amplification and barcode attachment. Libraries are then amplified by PCR and subsequently cleaned to prepare for hybridization.

Hybridization

Biotinylated oligonucleotides, SureSelect Human All Exon V5/V6 (Agilent Technologies, Santa Clara, CA, USA), that are specific to exon regions only are hybridized to the libraries in a 24 hour incubation at 65°C. Dynabeads MyOne Streptavidin T1 magnetic beads (ThermoFisher Scientific, Waltham, MA, USA) are used to separate exons from introns and other impurities. Streptavidin attaches to the biotin on the oligonucleotide-DNA hybrid and a magnet is used to separate them (Figure 16).

A final PCR is performed on all libraries using the same forward primer but with distinctive 8bp reverse primers. These act as barcodes to enable libraries to be distinguished from one another when pooled on the flow cell during sequencing. Molarity and quality were determined in all samples using the Agilent D1000 ScreenTape System (Agilent Technologies, Santa Clara, CA, USA) and a qPCR. The sample concentrations are then adjusted to 10nM and pooled together prior to be entered into the sequencer. Paired-end 125bp sequencing was performed on HiSeq 2500/4000 Illumina platform.

After sequencing the data is extracted and the libraries are cleaned up and separated by the bioinformatics team at the Functional Genomics Centre Zurich (FGCZ) according to the barcodes.

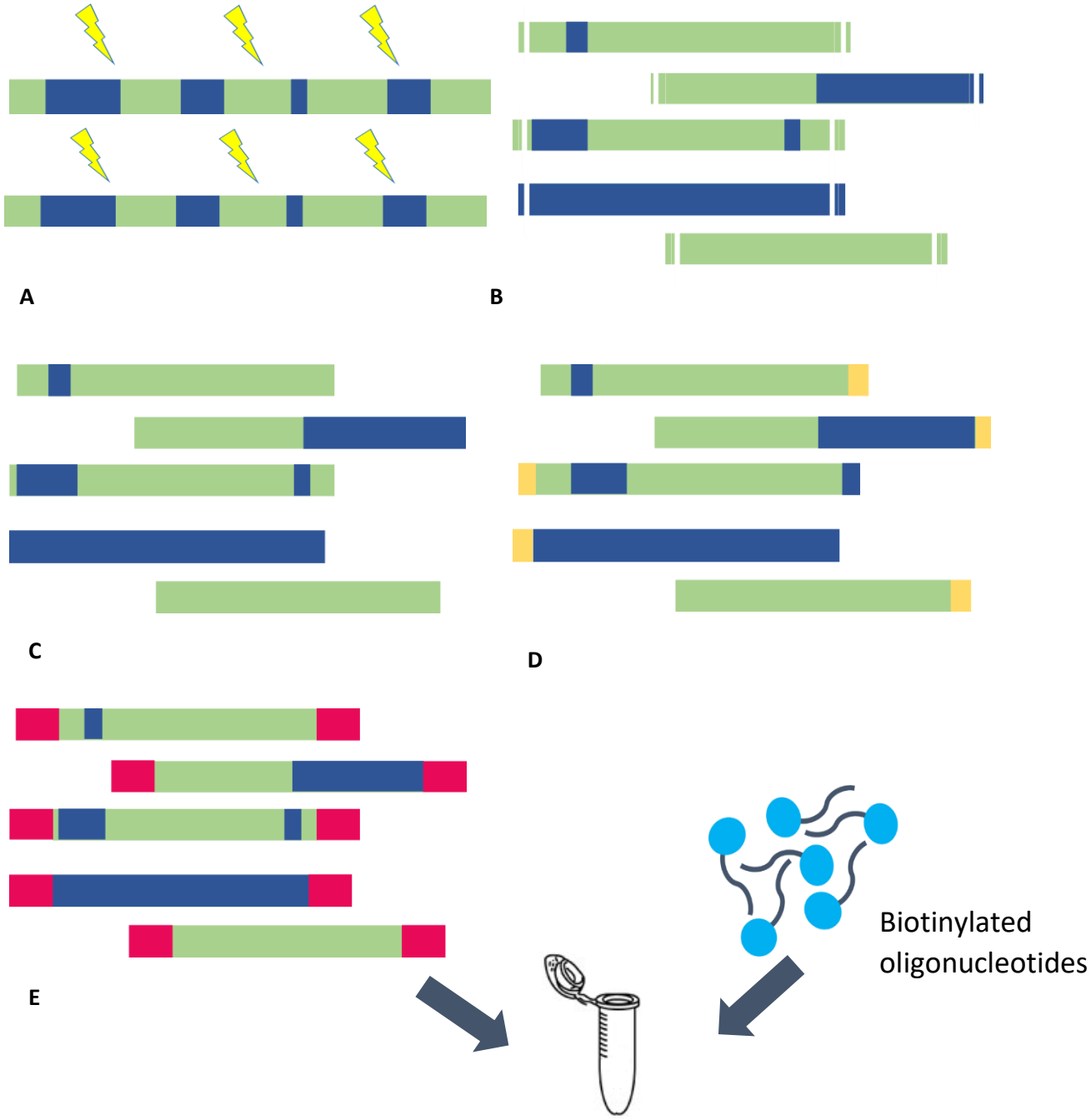


Figure 16.

A) Fragmentation of the genomic DNA, green is intronic regions and blue is exonic
B) Fragmented DNA has blunt ends
C) Blunt ends fixed by end repair
D) Adenylation (yellow) is performed on the 3 prime end.
E) Adaptor are ligated to the adenylated regions. The prepared library is bound to biotinylated oligonucleotides in a 24hr reaction.

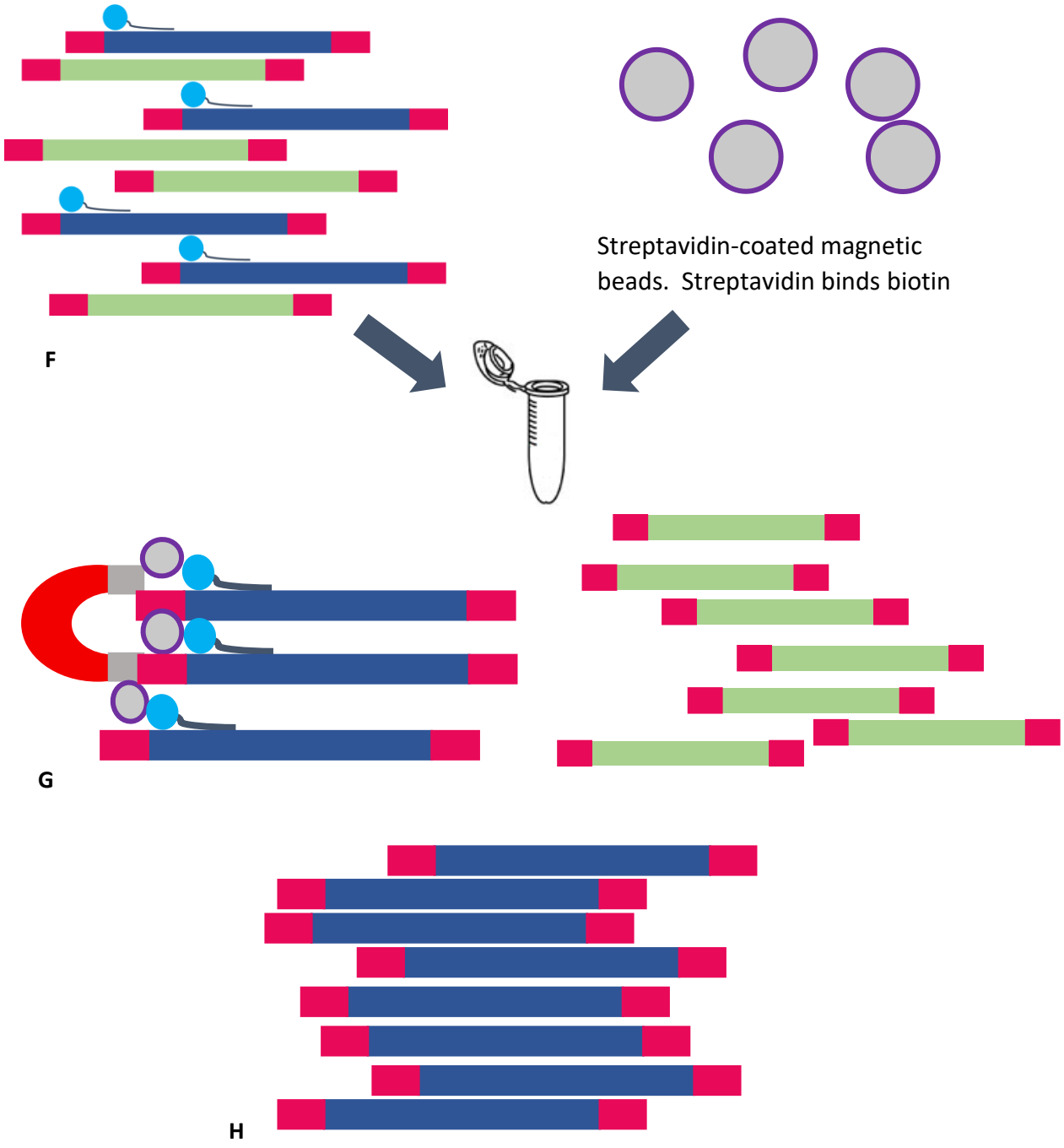


Figure 16.
F) Biotinylated oligonucleotides bound to the exons only. Streptavidin-coated magnetic beads added.
G) Magnetic beads coated with streptavidin binds to biotin on oligonucleotides. Magnet is used to separate exons from introns
H) Exons are amplified and ready for attachment to Illumina flow-cell

Bioinformatics Analysis

2. Genome analysis toolkit (GATK)

Sequencing output files were received in the FASTQ format. FASTQ files contain sequence data in text form with corresponding quality scores. Paired-end sequencing results in a forward and reverse file for each sample both averaging 2 to 4 GB each. FASTQ files were aligned to the reference genome (hg19) using Burrows-Wheeler Aligner (BWA-MEM) (SourceForge). BWA-MEM works by creating a compressed full-text substring index using Burrows-Wheeler transformation for compression. This can be used to detect patterns, their locations and the number of occurrences of these patterns in the text (SourceForge). The output of this is a SAM file which contains aligned sequence data in a tab delimited form (Li and Durbin 2010). This is then converted to a binary form called BAM file which, like a SAM file, lists the sequence data but is highly compressed, therefore the file size is smaller, making subsequent analysis steps faster.

Quality control and analysis were performed using the GATKv3.5 guidelines from Broad Institute using reference genome hg19 (McKenna, Hanna et al. 2010, DePristo, Banks et al. 2011, Auwera, Carneiro et al. 2013). Deduplication of sequencing data to remove reads in areas with abnormally high numbers was performed using Picard Tools v2.7.1. This is to ensure read statistics such as read depth are not skewed by PCR artefacts. BaseRecalibrator (GATK) was used to adjust base quality and detect errors in base quality. Output from the analysis pipeline is an altered BAM file.

Files were individually analysed for variations from the reference genome (hg19) using HaplotypeCaller from Broad Institute. HaplotypeCaller works by searching for regions in the

sample that differ from the genome and then reassembles reads at this region allowing highly accurate calling of SNPs and indels even is areas with more than one variant present (Highnam, Wang et al. 2015).

3. Somatic variant analysis

3 somatic callers, VarScan v2.3.9 (Koboldt, Zhang et al. 2012), MuTect 2 (Cibulskis, Lawrence et al. 2013) and Strelka Somatic Caller (Saunders, Wong et al. 2012) were used for each sample pair (Blood = Germline, Skin = Somatic). All somatic callers were originally developed to identify somatic mutations in cancer samples allowing for high sample heterogeneity by using a non-cancerous in-patient control.

VarScan2

VarScan2 uses exome data and searches for somatic mutations and copy number alterations by using a heuristics and a Fisher's exact test while analysing germline and somatic samples simultaneously (Koboldt, Zhang et al. 2012). Fisher's exact tests analyse frequency distribution data to find statistical significance. This method is often used when sample sizes are small. Heuristics is a method of problem solving which make assumptions that are not considered "perfect" but uses available data and methods to find the most optimal solution possible.

Strelka

Strelka detects small indels and SNPs while simultaneously analysing germline and somatic samples. It uses a Bayesian approach, which is a statistical method where the probability of the hypothesis is changed and updated according to new evidence or data available. The high level of sensitivity is achieved by considering the normal sample as germline plus some incidental noise and the somatic sample as germline plus some somatic variation (Saunders, Wong et al. 2012).

MuTect2

MuTect2 is similar to the previous version, MuTect, which uses a Bayesian classifier and highly specific filters to enable detection of somatic mutations at low allelic fractions. Unlike MuTect, MuTect2 uses *de novo* assembly of haplotypes in a region with detected variation (Cibulskis, Lawrence et al. 2013).

4. Annotation and quality control

Annotation of somatic mutations is performed by Combined Annotation Dependent Depletion (CADD v1.3) (Wang, Li et al. 2010, Koboldt, Zhang et al. 2012, Kircher, Witten et al. 2014). CADD reports findings from multiple databases (e.g. PolyPhen, SnpSift) showing all types of information on the mutations e.g. minor allele frequency (MAF), amino acid changes and positions and damage prediction. CADD also has a single metric which integrates and weighs information from other databases into one score which assess how "damaging" a mutation is. CADD does this by ranking the ~8.6 billion single nucleotide variations (SNV) according to their magnitude and then scales the Phred score from 0-40. Scores of >20 indicate the SNV is in the top 1% of damaging mutations and SNVs with >30 are in the top 0.01%.

All mutations of interest were finally quality controlled manually using visualization software. Integrative Genomics Viewer (IGV) (Thorvaldsdóttir, Robinson et al. 2013) was used to compare sequencing data from the analysis pipeline to the reference genome. IGV is an application, which allows for interactive exploration of your data. The input file format used here is BAM however it can support other file types. IGV allows you to choose your reference genome and analyse your genome or other NGS data in this context. It uses a process called "data-tiling" where summarizations of the data are precomputed at several resolution scales. Each scale is separated into tiles along the genome that are viewable on the display. It is possible to upload several exomes at a time therefore the sample of interest can also be compared inter- and intra-patient to confirm true somatic mutations and confirm rarity. To do this the somatic sample is compared to the corresponding germline sample as well as uninvolved control exomes, to confirm a mutation of interest is absent from all controls.

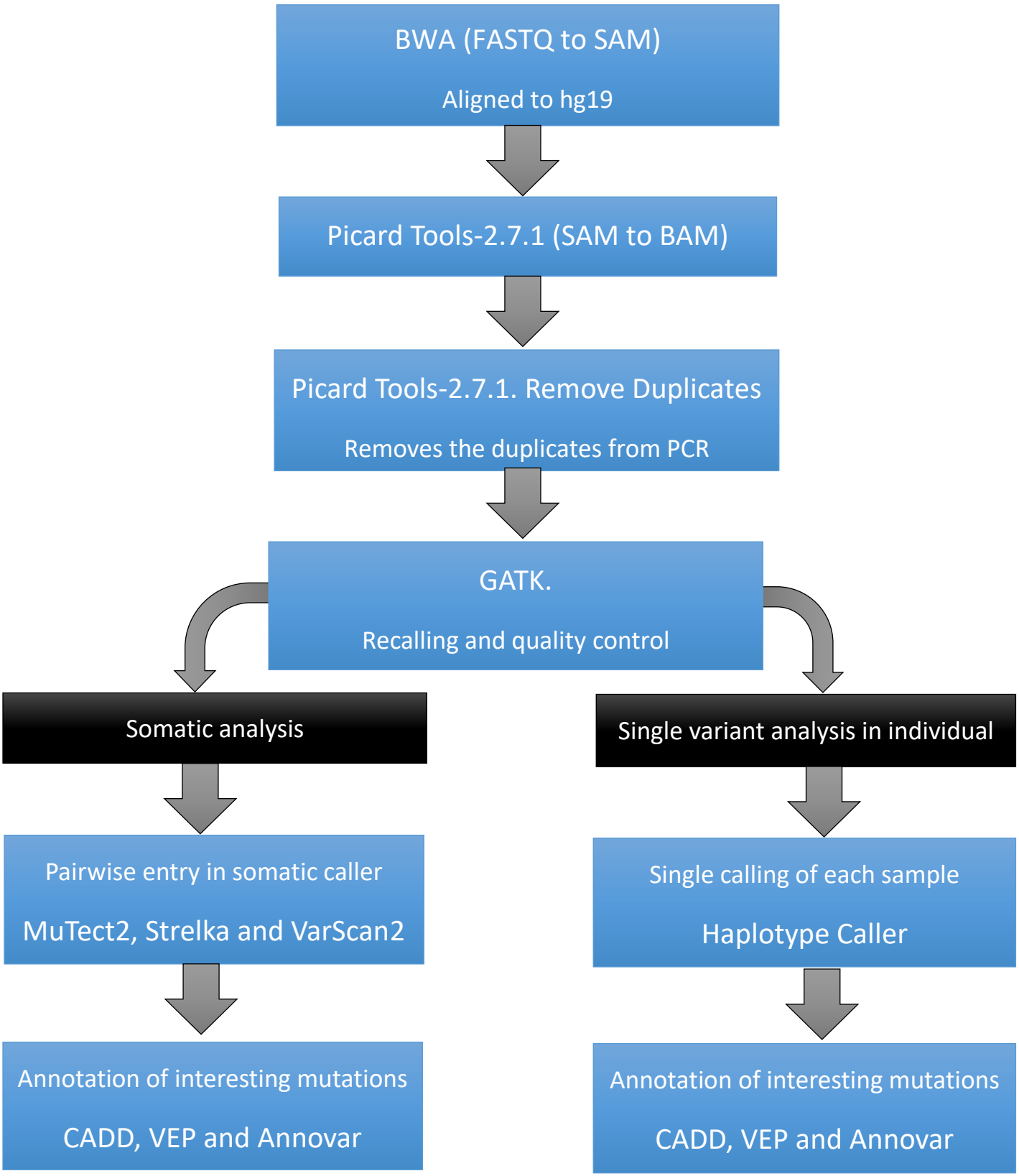


Figure 17. Bioinformatic analysis. Step by step analysis of WES data. Each sample is analysed together and individually.

5. Germline analysis

In order to perform an analysis of germline mutations 80 in-house samples from the Caucasian cluster without LLS were chosen as controls using principal component analysis. A multisample genotype-called VCF file for controls (n = 80) and cases (n = 19) was created using GenotypeGVCFs from GATK v3.6. A germline analysis was performed using Efficient and Parallelizable Association Container Toolbox (EPACTS) (Kang).

EPACTS performs two types of analysis. First, it searches for a single variant that is found in the affected samples and controls with different allele frequencies. Second, it measures the burden of mutations in genes and searches for a gene that has an increase of rare variants in cases when compared to controls.

EPACTS also filters samples according to their consequence: truncating, alternative and synonymous. Truncating mutations result in a premature stop codon, alternative mutation result in a change in amino acid and synonymous mutations are when the nucleotide changes but this still codes for the same amino acid. Mutations are further separated into dominant and recessive mutations. Dominant mutations are when there is a single copy of the mutation. A recessive mutation is when both copies of the gene carry the mutation.

To search for a single variant, EPACTS performed a Wald test. A Wald test is a parametric statistical analysis which eliminates variants that are likely having no effect. For a gene-based approach, Fisher's exact and SKAT-O tests were used. A SKAT-O test is a test that looks at the multiple-variants at the same region and looks for association with a phenotype (Wu, Lee et al. 2011).

Output data was separated according to dominant (MAF < 0.01) and recessive (MAF < 0.05) mutations. The MAF for recessive mutations have a higher cut off. This is because the MAF deals with the frequency of a mutation in the population where mutations are predominantly in one copy of the gene. Two copies of the mutation is much less likely however, since MAF does not take this into account, the stringency setting must be more lenient.

For each output Q-Q plots and Manhattan plots were created. The QQ plots represent the expected p-vales of the data plotted against the actual p-values. Output Manhattan plots represent p-values of each data point according to chromosome.

6. HLA analysis of LLS patients

Three HLA allotype callers were used to determine the HLA allotypes of the LLS patients and 60 controls that do not have LLS. The callers use NGS data to infer allotypes. The HLA gene cluster is located on chromosome 6. HLA typing has previously been time consuming as it required PCR amplification with specific primers or the use of sequence-specific oligonucleotides during WES or NGS. Several callers have been developed to find HLA allotypes of patients without the need of these techniques (Szolek, Schubert et al. 2014). As NGS technologies such as WES and WGS are adopted frequently in clinics now, it is possible to use these callers on the data already available and avoid repeat experiments. The callers used in this project are OptiType (Szolek, Schubert et al. 2014), VBSeq (Nariai, Kojima et al. 2015) and BWA-kit (Burrows-Wheeler-Aligner-kit 2016).

OptiType

This program uses NGS data to infer HLA allotypes using an integer linear programming approach. It has been shown to achieve an accuracy of 97% (Szolek, Schubert et al. 2014).

VBSeq

WES data is realigned along the reference genome (hg19) using BWA-MEM and reads that mapped along the HLA regions are extracted. Unmapped sequences are then removed and re-aligned along all HLA sequences from the IMGT/HLA database and then extracted. Both extracted files were combined and aligned with BWA-MEM. HLA alleles are estimated using Bayesian inference (Nariai, Kojima et al. 2015).

BWA-kit

BWA-kit is a package for HLA allotyping. It has two scripts. One for downloading the human reference sequence and another for mapping reads to the HLA regions. It can also use alternative contiguous sequences which represent different haplotypes, to improve read-mapping.

Inconsistencies were found in MHC Class II between all 3 HLA callers in the same patient, to this end the analysis just focuses on MHC Class I (Major, Rigo et al. 2013).

7. Comparative genomic hybridization

High quality DNA samples of 3 patients (i.e. samples with the lowest level of degradation) were chosen for Comparative Genomic Hybridization (CGH) (Beaudet 2013, Asadollahi, Oneda et al. 2014). The array was performed using CytoScan (2.6 million markers) (Affymetrix Inc., Santa Clara, CA, USA) to find chromosomal structural abnormalities such as large-scale deletions and duplications. The patient karyotypes were then compared to 820 in-house controls (Institute of Medical Genetics, Zürich) and 1038 Affymetrix controls with healthy karyotypes. The data set of each patient's sample was evaluated with Affymetrix® Chromosome Analysis Suite (ChAS). This is used to rule out commonly found or non-damaging chromosomal copy number abnormalities and chromosomal aberrations.

8. Copy number variation (CNV) analysis

Copy number variation (CNV) is an event that differ between two genomes that is larger than 1 kilobase. These changes are too small to be seen in genomic hybridization experiments as these are not sensitive enough to detect such small alterations. CNVs encompasses gains and losses of DNA. Somatic CNV events such as duplication of oncogenes and deletion of tumour suppressors have been described in cancer (Liu, Morrison et al. 2013, Qin, Liu et al. 2015). CNV events have not, as of yet, been implicated in non-cancer diseases.

CNV analysis was performed using Sequenza v2.1.2 (Favero, Joshi et al. 2015) using WES data. Pileup files were generated from exome BAM files using Samtools v1.3.1 (Li, Handsaker et al. 2009). Pileup files contain base-pair information facilitating the calling of SNPs and indels. Pairwise analysis of samples was performed using pileup2seqz from the Sequenza package (Figure 18) (Favero, Joshi et al. 2015).

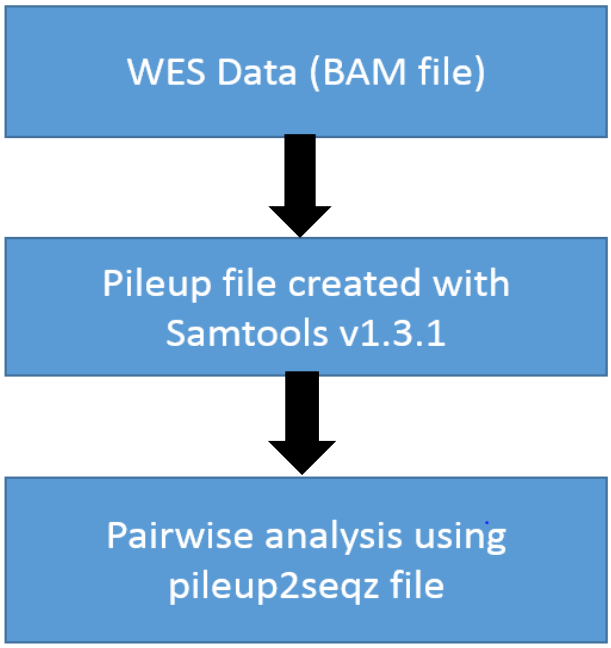


Figure 18. Sequenza analysis of CNV

9. Loss of heterozygosity

Samtools-v1.3.1 (Li, Handsaker et al. 2009) was used for input file preparation for the somatic mutation callers. VarScan v2.3.9 (Koboldt, Zhang et al. 2012) was used to investigate possible loss of heterozygosity (LOH) mutations as a genetic cause in LLS. LOH is an event where a region of DNA is lost on one chromosome resulting in only one copy of a gene. This is damaging as a heterozygous mutation becomes homozygous due to the lack of a second copy of the gene.

10. Determination of the sensitivity of somatic callers and deep sequencing

A test BAM file containing chromosome 21 only was created for 5 test samples using PrintReads (GATK). BAMSurgeon (Ewing, Houlahan et al. 2015) was used to insert a directed synthetic mutation in the test BAM files. BAMSurgeon can insert mutations at a designated allelic fraction. Mutations were inserted at 2 regions with different read depths (DP), DP20 and DP100. DP is how many times a single nucleotide has been sequenced. These fabricated mutations were inserted at allelic fractions ranging from 0.05 to 1 at 0.05 increments. These files were then entered into all 3 somatic callers to determine the degree of coverage necessary for a mutation to be called in a highly heterogeneous sample. The control samples were the same chromosome 21 test files without the inserted mutation.

Results

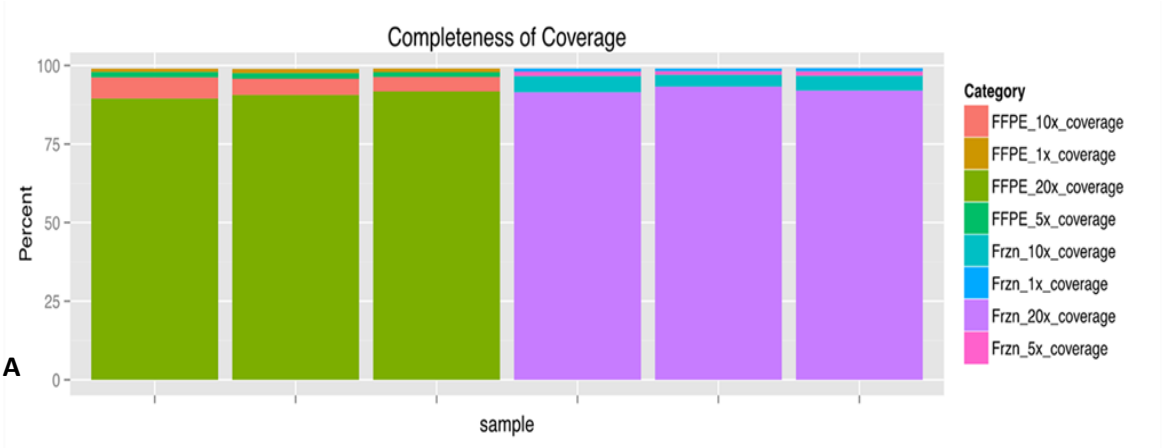
Analysis of sequencing data

1. Sample type comparison

Quantification of differences between preservation types was performed to confirm conformity and legitimise inclusion of both FFPE and SF samples into one study. Skin of 3 healthy persons was received post-routine surgery, split in half with half placed in formalin and half snap frozen. An analysis of the coverage and quality of FFPE samples when compared to SF of the same patient was performed. Sample types were analysed and compared in various aspects including average coverage and specificity. Average coverage for SF (DP 24.52 ± 2.6, n=3) and FFPE (DP 26.21 ± 1.7, n=3) (Table 3) are comparable with >90% of exons achieving over 20x coverage (Figure 19A). Specificity of SF (57.01 ± 0.47, n=3) and FFPE (57.67 ± 1.98, n=3) (Table 3) are also comparable. Number of C to T mutations was determined in the FFPE samples and compared to the SF skin of the same patients. Burden of C changes in FFPE was found to be comparable to that of SF samples (Figure 19B). No significant changes were found between sample types, so we concluded both FFPE and SF could be included in the same study.

	Average coverage	Specificity
SF (n=3)	24 ± 3	57 ± 0.47
FFPE (n=3)	26 ± 2	58 ± 2

Table 3. Quality of WES data of FFPE compared to SF



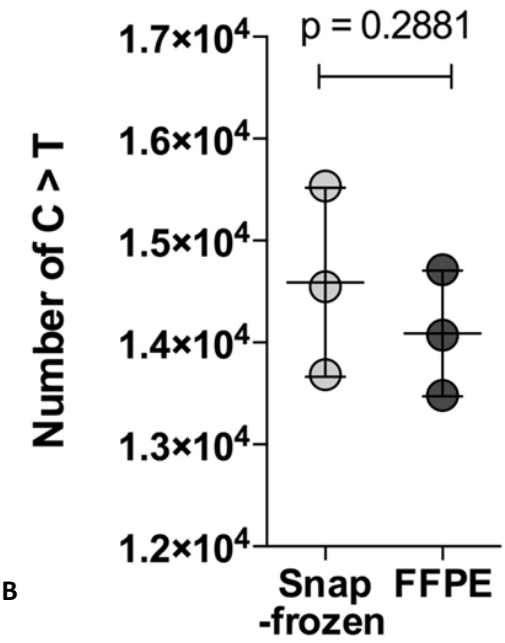


Figure 19. Comparison of modes of skin preservation.
A) Coverage of FFPE vs SF in WES data,
B) Number of C to T mutations in corresponding samples

2. Somatic analysis of LLS patients

After sequencing, the analysis pipeline was followed and output data underwent quality control analysis. The analysis showed average read depth (DP) of >80 was achieved for each sample. The CADD database was used to annotate the somatic variants called by the 3 variant programs. Somatic analysis of annotated mutations began with the search for a rare and predicted damaging mutation in common between patients. Mutations with a minor allele frequency (MAF) over 0.01 in European populations were filtered out. MAF frequencies were determined from data of the Thousand Genome Project (TG_MAF) and Exome Sequencing Project (ESP_MAF). Mutations with a CADD score below 15 were also filtered out to isolate mutations that were predicted to be damaging. Synonymous mutations were removed. Results from PolyPhen and SnpSift were consulted when scrutinising each mutation. No single damaging and rare SNP mutation was found to be in common between more than 5 patients (Strelka) (Table 4). Use of Integrative Genome Viewer (IGV) to visualise these mutations ruled them out as promising candidates because of presence of mutations in unaffected controls or presence of the mutation in germline samples. Sometimes the DP of the mutation in the germline is too low for detection by the program therefore it will conclude it is a somatic mutation. No indel was found in common between more than 2 patients (Table 4).

Numbers of patients	SNPs	Indels
1	60’681	3’840
2	5’392	340
3	118	0
4	11	0
5	1	0
6... 19	0	0

Table 4. Number of SNPs and indels found in common between patients, non-synonymous only

3. Gene-based somatic analysis

A similar analysis was performed with the aim of finding a single gene affected with different somatic, rare and damaging SNP and indel mutations in common between the patients. This was achieved by filtering for the rare and damaging somatic mutations of all patients. Genes affected with these somatic mutations were then isolated. The list of affected genes of all patients were compiled. The number that each gene appears in the compiled list was found. Genes that were most frequently affected were analysed first.

No single gene burdened with somatic mutations was found in common between the patients. 10 patients appeared to have *HLA-DRB1* mutations, however IGV visualization revealed that these mutations were most likely due to aligning artefacts as HLA regions have high variation from the reference genome. 5 patients had a SNP in *PRSS3* and *CDC27* genes and 4 had a SNP mutation in *ANKRD36* and *MUC16* genes (Table 5). Both genes were ruled out as increased scrutiny into these somatic mutations showed these mutations were not true somatic mutations as some were present in the control samples in a low allelic fraction undetectable by the somatic mutation callers, making them appear somatic to the program. In addition, several mutations were in fact germline upon visualization.

Number of samples	Number of affected genes
1	595
2	35
3	11
4	2
5	2
6... 19	0

Table 5. Somatically affected genes, rare, damaging (synonymous included)

4. Pathway-based analysis

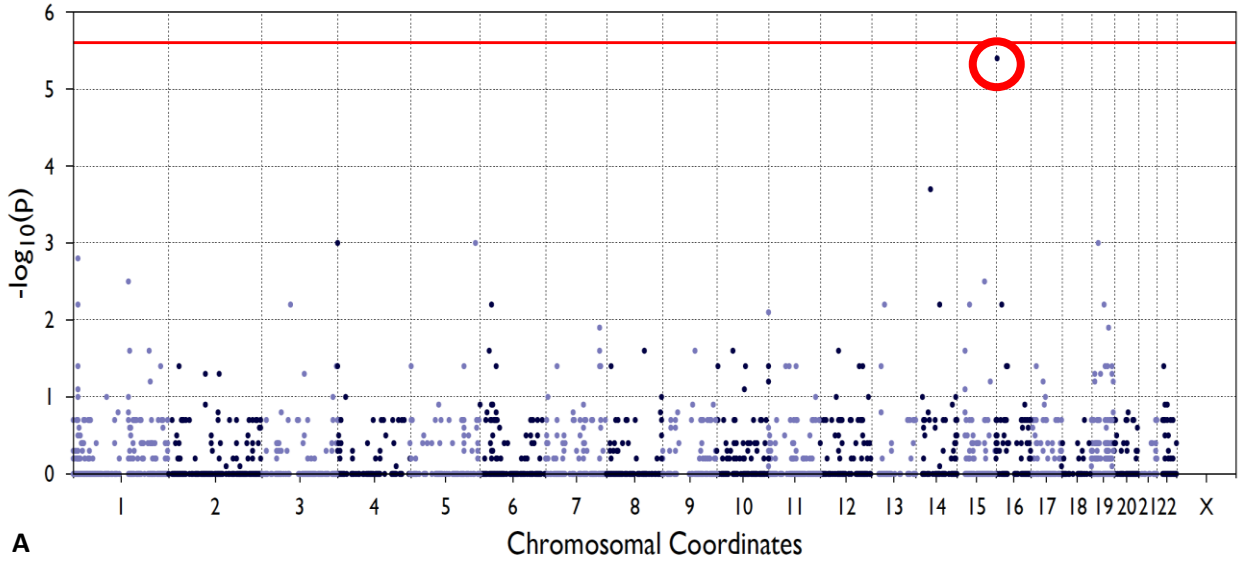
STRING online database v10.5 was used to find associations between somatically affected proteins in the LLS cohort with the hopes of unearthing a particular pathway with a high burden of mutations. All genes affected with a somatic mutation in any patient were entered into the program. Pathways with the highest amount of mutations were isolated and the mutations involved were individually analysed. Proteins in the ubiquitin mediated proteolysis (UMP) pathway were found to have several somatic mutations. All involved somatic mutations were shown to be false positives using IGV as the mutations in question were actually germline and/or found in control samples.

5. Germline analysis

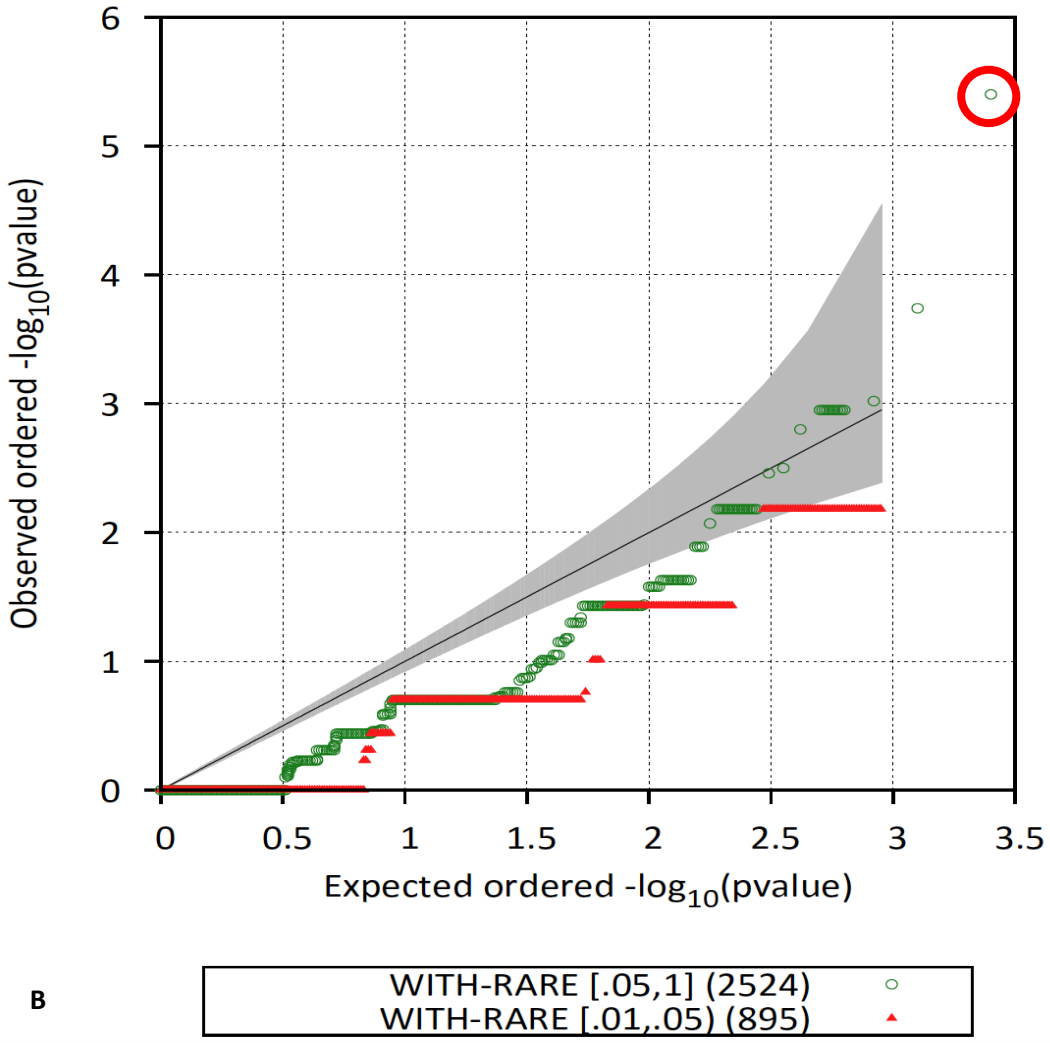
Gene-based germline analysis did not yield any gene burdened with mutations that were not present in controls.

There was one gene, *DNASE1L2*, affected with a near exome-wide significant p-value (3.9^{-6}). This gene was affected with 3 recessive mutations, 2 non-synonymous and 1 synonymous (16:2287496_T/C, 16:2287501_T/C, 16:2287491_G/C) mutations (Figure 20). *DNASE1L2* is a paralog of *DNASE1*, an enzyme involved in binding of calcium ions and has deoxyribonuclease activity. The two, non-synonymous mutations had a MAF of 0.04 and the synonymous mutation had a MAF of 0.026. Use of IGV confirmed at least one of these mutations was present in every patient. However, more than one of these mutations was found in control samples (unrelated individuals) resulting in them being ruled out as risk factors for LLS.

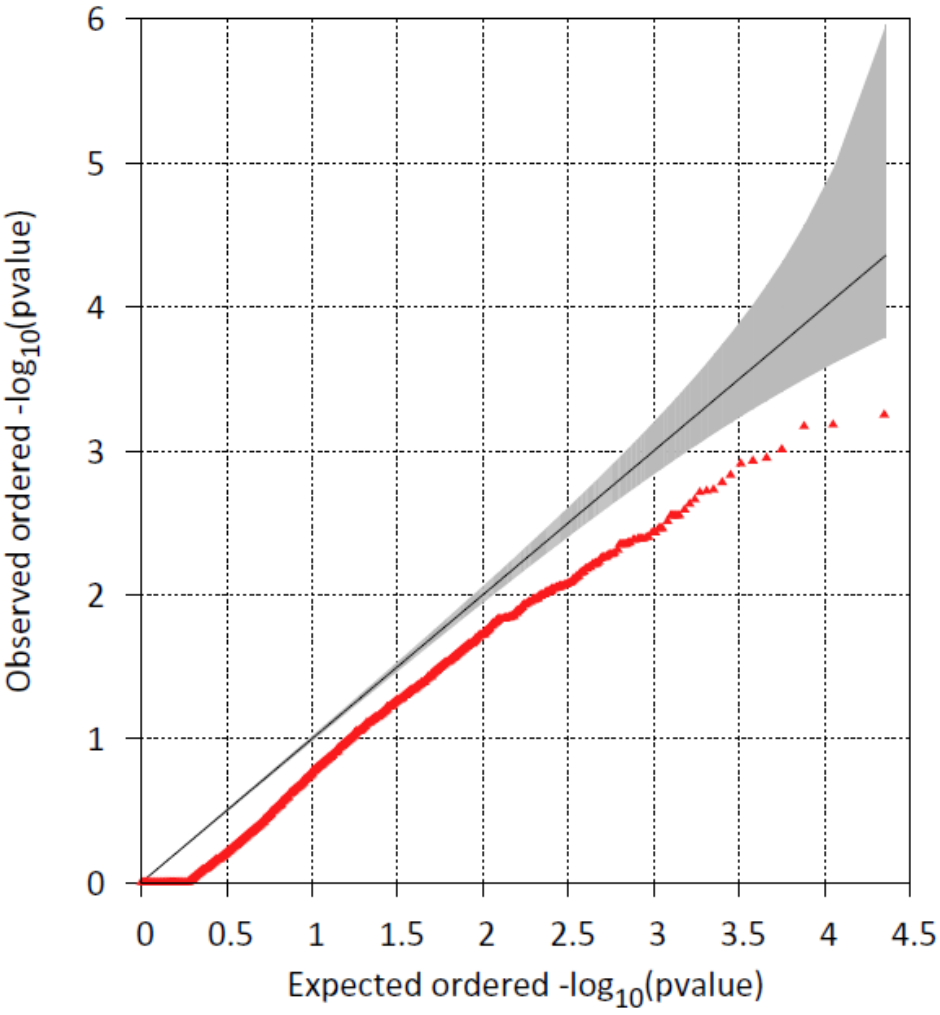
In single mutation analysis no p-value was found to have a higher than the expected threshold p-value (Figure 20C) however all mutations with a low p-value were visualized on IGV. A heterozygous mutation (rs74700936) was found in one gene *ARNT2,* aryl hydrocarbon receptor nuclear translocator, it is involved in the xenobiotic responsive element. This mutation was found in in 5 of the 19 patients. The mutation was found to be much less common in controls. This is a mutation in the 3' UTR (untranslated region) in exon 19. Previous mutations in exon 13 of this gene have been described in Webb-Dattani syndrome, a neurological disorder. EPACTS predicted it as significant (p-value 0.0083616). However, it was ruled out as a causative mutation of LLS as annotation program CADD predicted it to have very low effect on protein function (CADD=1.91), SnpSift and PolyPhen could not predict an effect. Its MAF was quite low at 0.04 however, this is not a very rare mutation.


-log10(P)
0
1
2
3
4
5
6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Chromosomal Coordinates

A


Observed ordered -log10(pvalue)
Expected ordered -log10(pvalue)
0
0.5
1
1.5
2
2.5
3
3.5
0
1
2
3
4
5
6
WITH-RARE [.05,1] (2524)
WITH-RARE [.01,.05) (895)

B



C

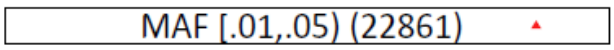


Figure 20. A) Manhattan plot of the p-values of each gene being significantly affected with germline mutations plotted according to chromosome. Red ring highlights the gene *DNASE1L2*.
B) Is a QQ-plot of the expected p-values of genes vs the observed p-values. Red ring highlights the gene *DNASE1L2*.
C) QQ-plot of p-values of single variants.

Other somatic analyses

6. Comparative genomic hybridization

CGH analysis of 6 affected skin samples found no interesting chromosomal aberrations in common between patients. All aberrations in the patients were shown to be present in the control samples. All controls are healthy with normal karyotypes. Likewise, when patients were visualized individually, no significant or rare chromosomal aberrations was found. All aberrations were shown to be present in healthy control samples (Figure 21).

7. Copy number variation analysis

Analysis of somatic CNV events in all patients did not reveal an event in common between all patients. When comparing all the samples pairwise no single somatic CNV event was found in common between all patients. Indeed there was no event in common between any two of the patients. A second analysis focused on regions with different CNV events overlapping between patients. No region was affected with different CNV events in overlapping regions.

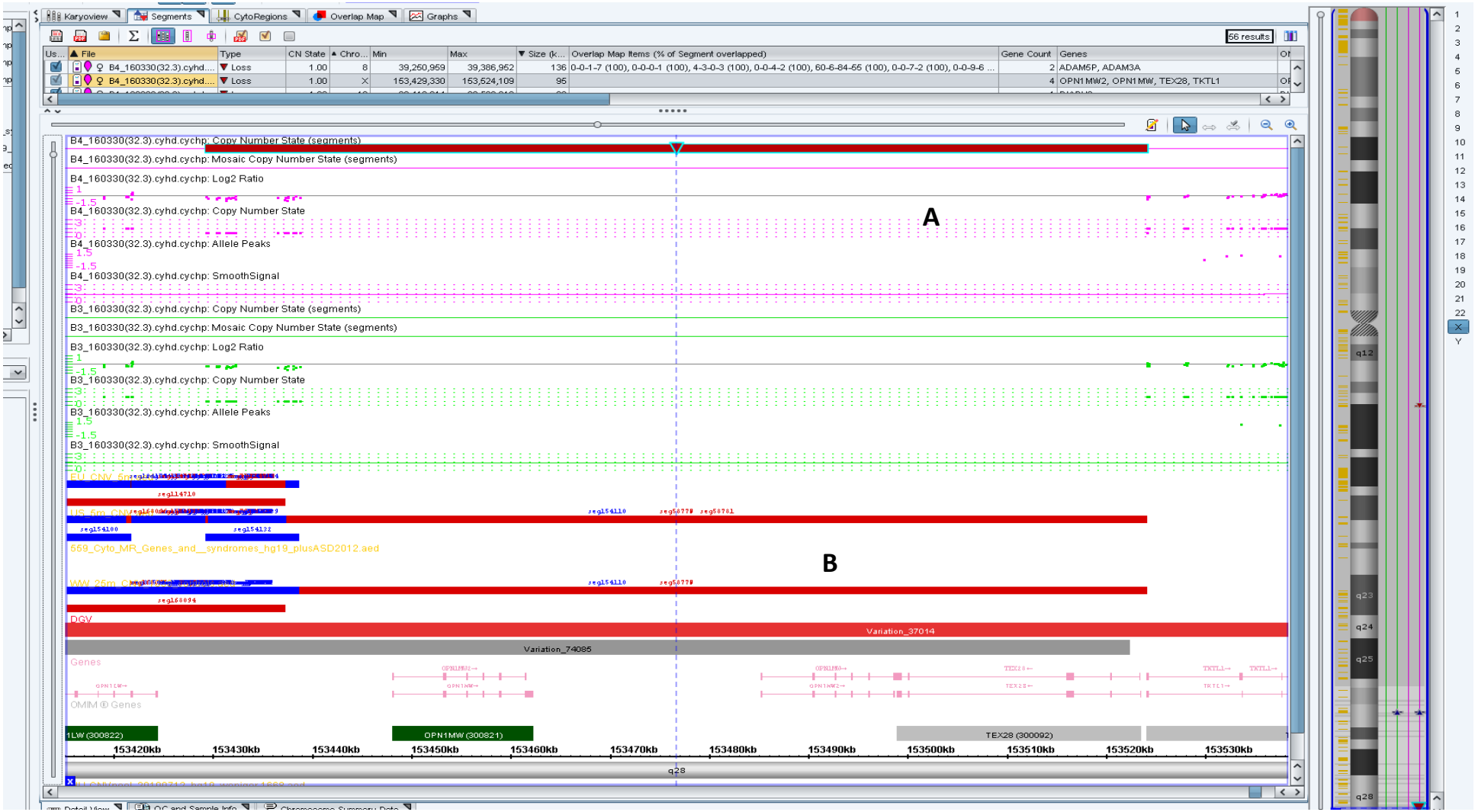


Figure 21. Screenshot of the ChAS program for viewing karyotypic data. Here the program shows A) a large scale deletion on chromosome 8 of a patient. B) Shows the same deletion and similar deletions affecting the same areas in controls.

8. HLA allotypes

HLA Class I analysis was performed on all patients. Results of MHC Class I allotypes were compared between the 3 allotype calling programs. Allotypes in all 3 were consistent in all patients between all 3 programs. MHC Class II genes, were inconsistent between all three callers and the results were unreliable with little to no overlap. It is difficult to call MHC Class II data from WES data and this region is highly variable and leads to difficulties for the program when calling. Therefore we had to restrict the analysis to focus on MHC Class I only. Allotypes of 50 confirmed Swiss exomes to act as controls.

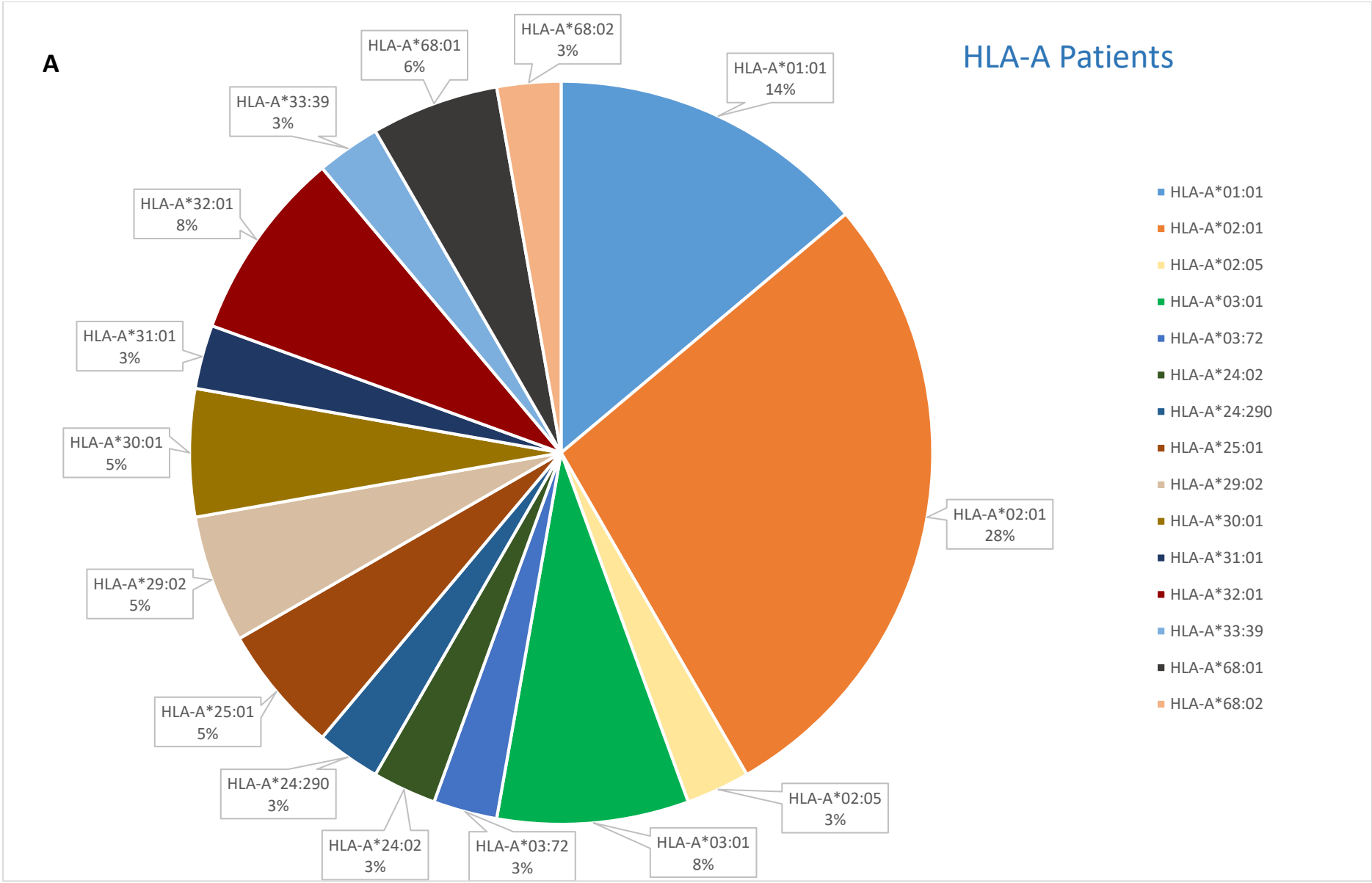
First, to evaluate the HLA allotypes of our patients the allotype of blood and skin were compared in each patient. A somatic alteration in allotype has not been previously described as a cause of genetic disease. It was concluded that there was no allotypic differences in the affected area. Second, the burden of certain allotypes was investigated within the cohort and compared to our controls. As LLS is an immunological disorder, a unifying allotype possibly acting as a driver of the disease was theorized. There was no allotype with a particularly high burden in the cohort. Third, our cohort was compared to published statistics from the Swiss population to hopefully unearth an unusual pattern (Buhler, Nunes et al. 2012). The samples in our cohort reflected the available data on Swiss allotypes and the HLA allotype was ruled out as a possible driver of LLS (Figure 22).

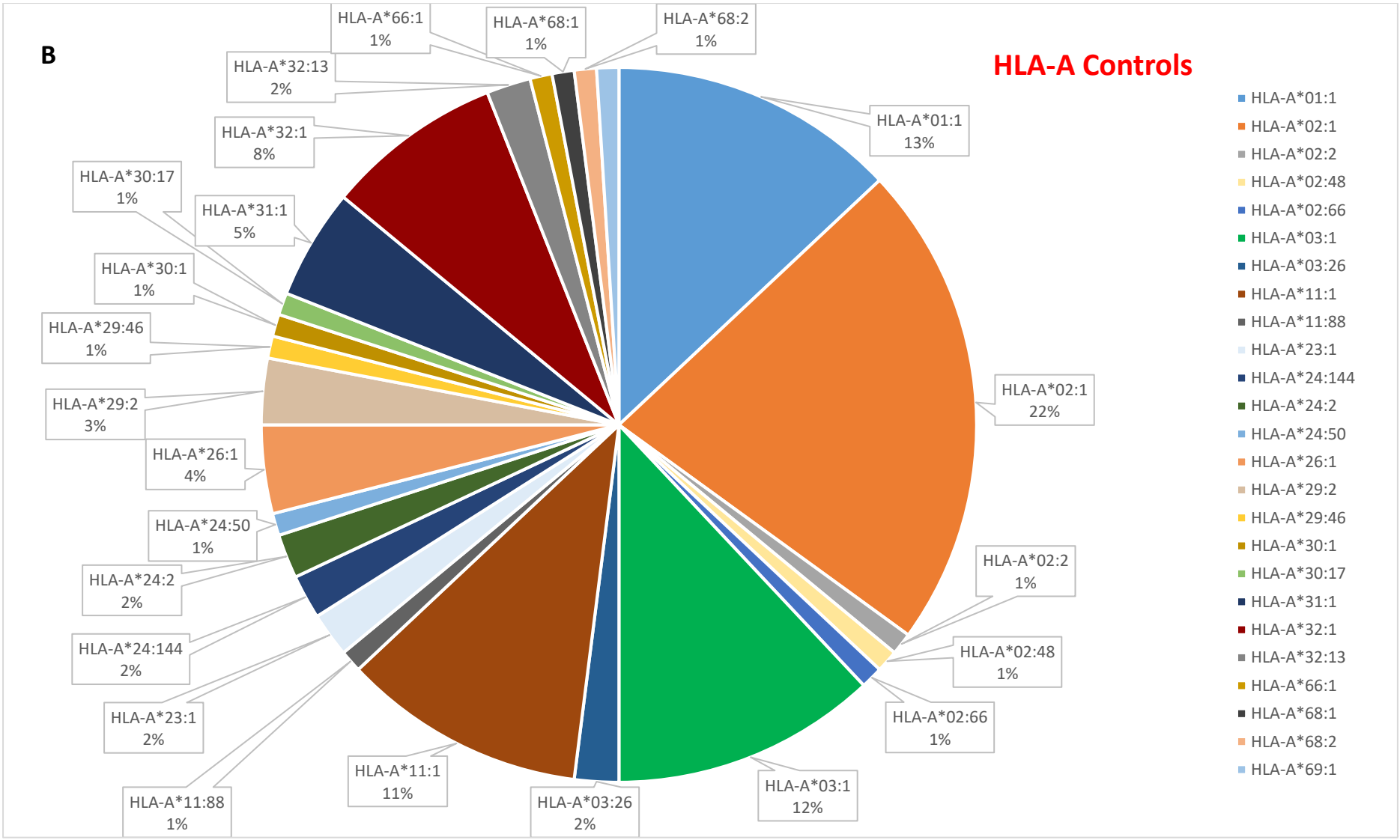
Of the HLA-A allotypes 02:01 is most common, 22% of the Swiss population has this allotype. This reflects our cohort where 28% have the 02:01 allotype. 14% of the Swiss population and 13% of the cohort have 01:01. 12% of patients and 8% of Swiss controls have 03:01 (Figure 22A, 22B). These are the 3 most common allotypes in the cohort and they roughly reflect the Swiss population.

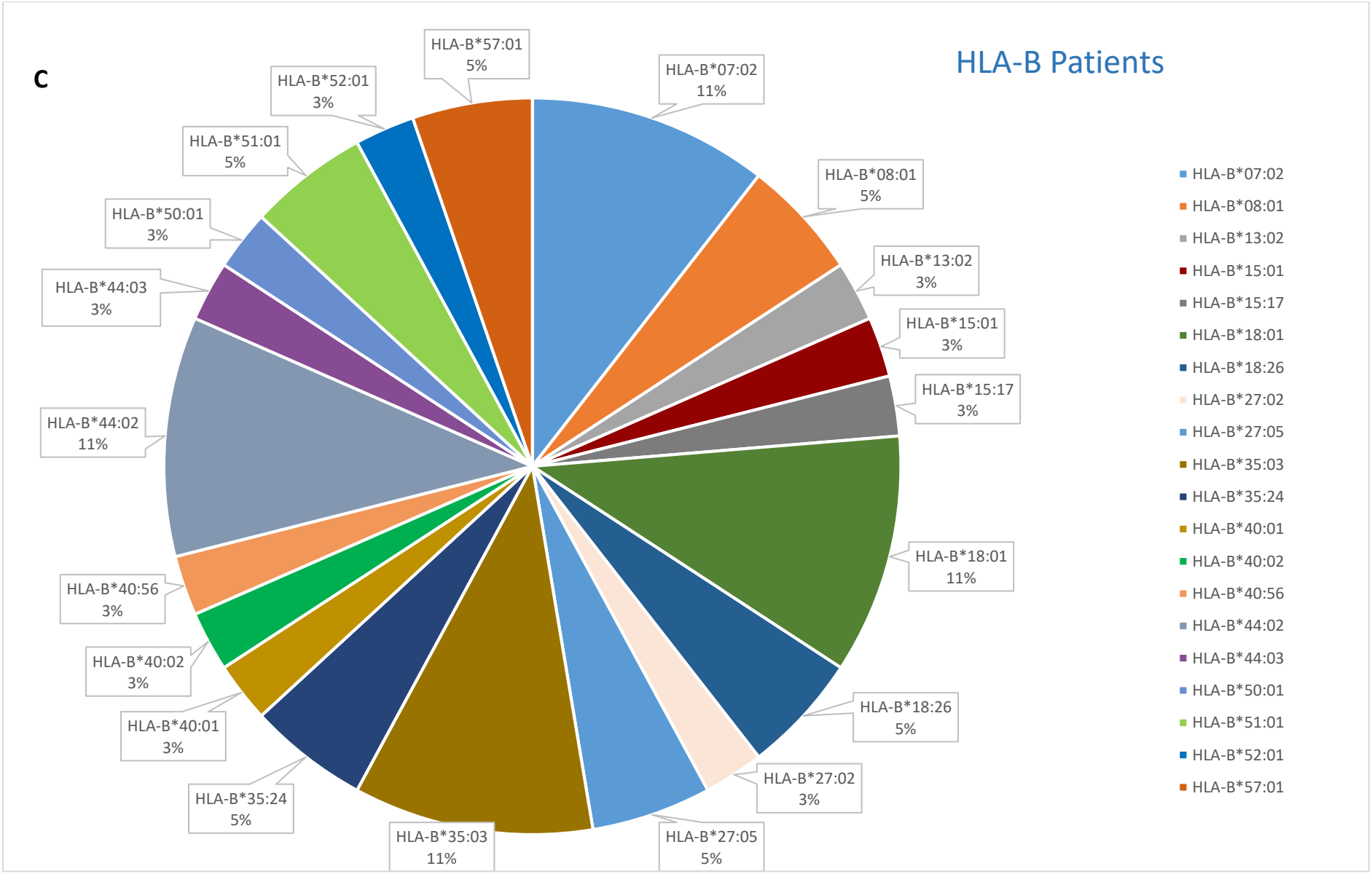
The most common HLA-B in Switzerland is 44:02 (11%) which reflects this cohort (11%). There is an abnormally high subtype found in the cohort, 35:03 is found in 11% of the cohort (Figure 22C, 22D). This does not reflect literature where it is rare ~1%. However, similar allotypes are more common, 35:01, 35:02 and 35:08 (6%, 2% and 1%). No more than 4 patients had this allotype so it doesn't implicate this as a factor.

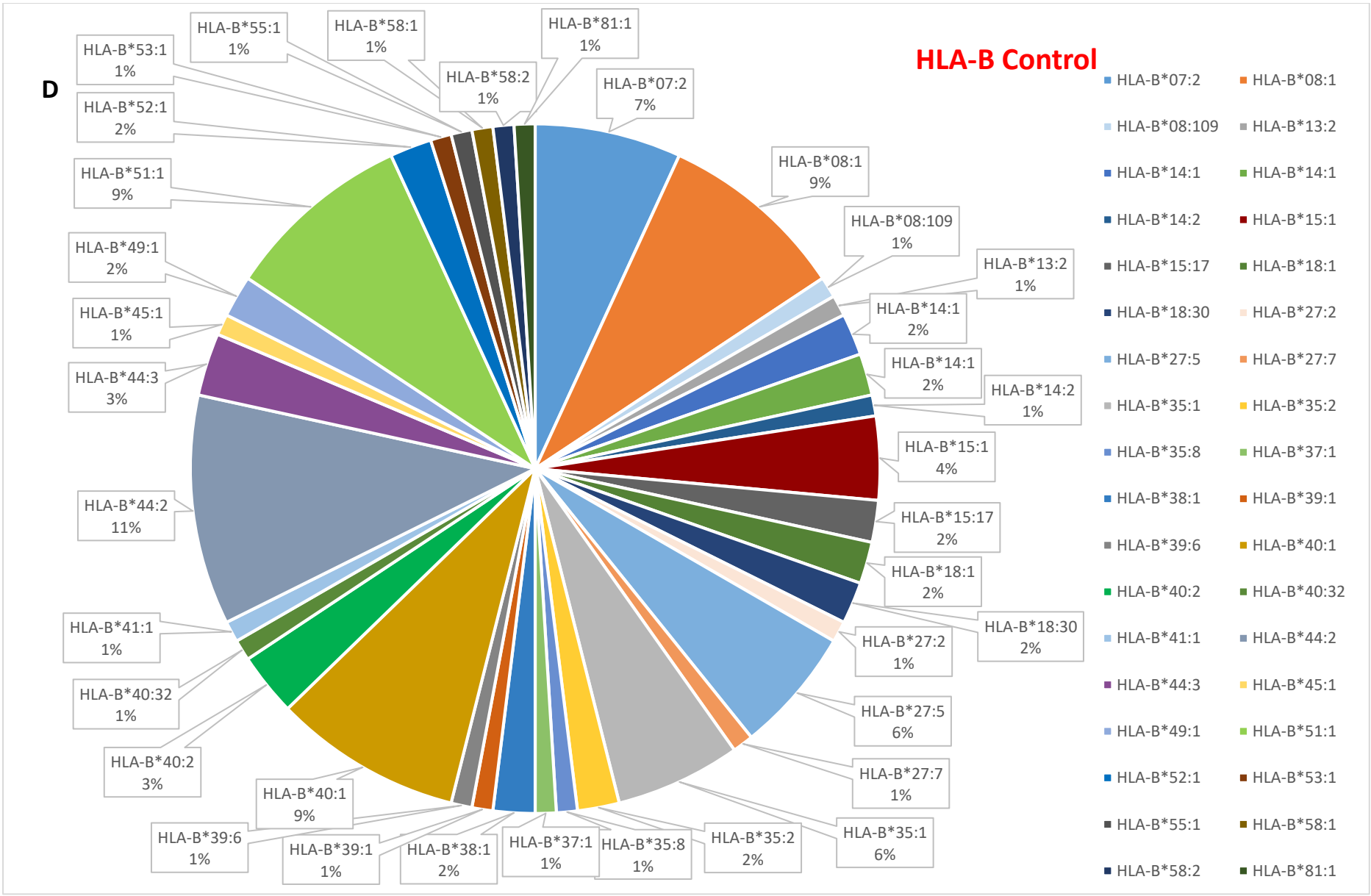
The most common HLA-C allotype in Switzerland is 07:01 (12%), however in this cohort only 5% of patients had this allotype. In the cohort 06:02 is the most common allotype (16%) but is less common in the Swiss population (6%). This is abnormally high because one patient had two copies of this allotype. 02:02 is also elevated, 13% in patients but only 7% in the Swiss population (Figure 22E, 22F).

Some allotypes in HLA-B and -C are elevated or decreased when comparing the cohort to the Swiss population. However, none of these allotypes were present in more than 6 of the patients. The abnormally high burden of these allotypes occurs because in some patients the allotype was present on both alleles causing the percentage affected to look abnormally high but not widespread in the cohort.

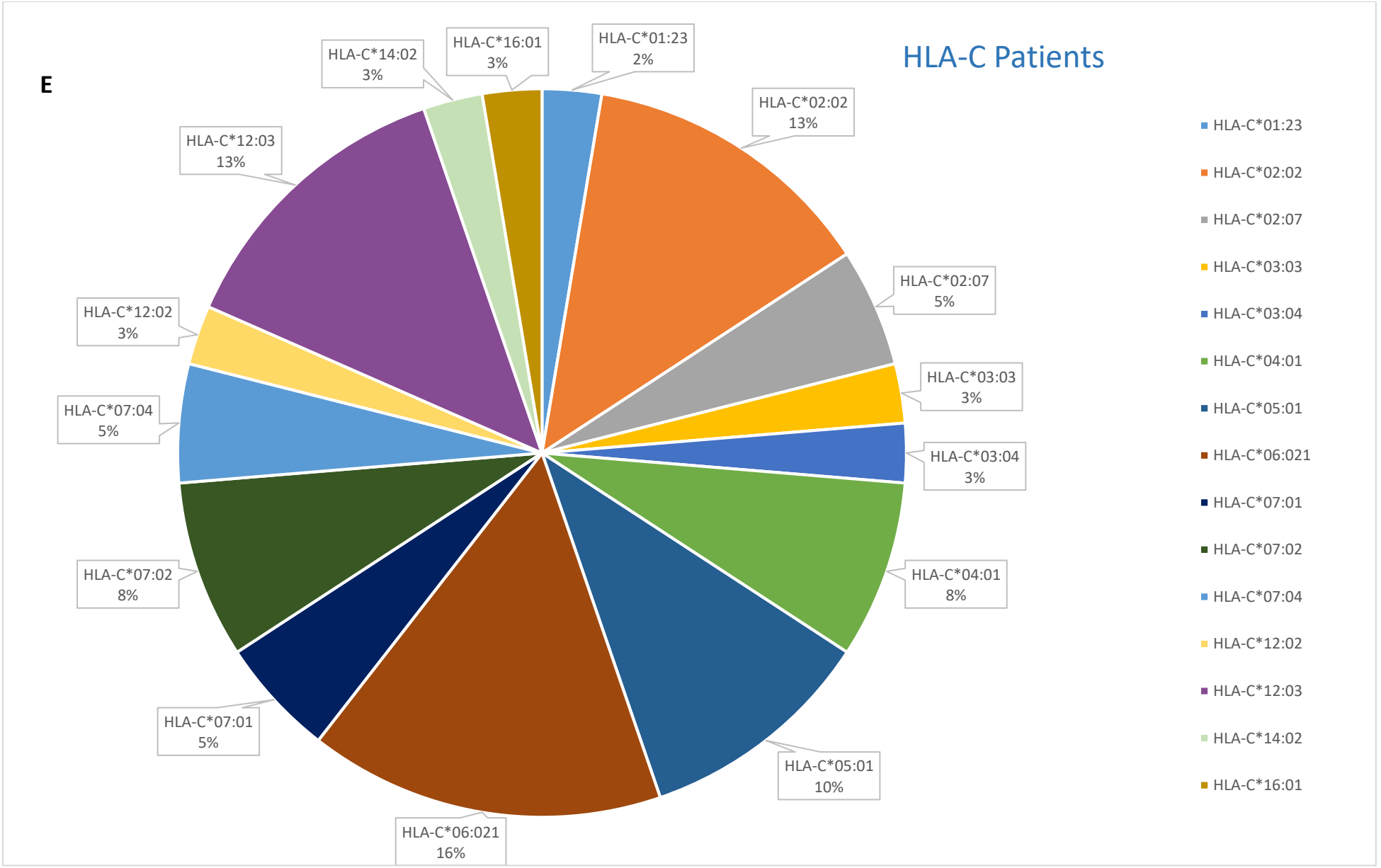

A
HLA-A Patients
HLA-A*01:01 14%
HLA-A*02:01 28%
HLA-A*02:05 3%
HLA-A*03:01 8%
HLA-A*03:72 3%
HLA-A*24:02 3%
HLA-A*24:290 3%
HLA-A*25:01 5%
HLA-A*29:02 5%
HLA-A*30:01 5%
HLA-A*31:01 3%
HLA-A*32:01 8%
HLA-A*33:39 3%
HLA-A*68:01 6%
HLA-A*68:02 3%
HLA-A*01:01
HLA-A*02:01
HLA-A*02:05
HLA-A*03:01
HLA-A*03:72
HLA-A*24:02
HLA-A*24:290
HLA-A*25:01
HLA-A*29:02
HLA-A*30:01
HLA-A*31:01
HLA-A*32:01
HLA-A*33:39
HLA-A*68:01
HLA-A*68:02


B
HLA-A Controls
HLA-A*01:1 13%
HLA-A*02:1 22%
HLA-A*02:2 1%
HLA-A*02:48 1%
HLA-A*02:66 1%
HLA-A*03:1 12%
HLA-A*03:26 2%
HLA-A*11:1 11%
HLA-A*11:88 1%
HLA-A*23:1 2%
HLA-A*24:144 2%
HLA-A*24:2 2%
HLA-A*24:50 1%
HLA-A*26:1 4%
HLA-A*29:2 3%
HLA-A*29:46 1%
HLA-A*30:1 1%
HLA-A*30:17 1%
HLA-A*31:1 5%
HLA-A*32:1 8%
HLA-A*32:13 2%
HLA-A*66:1 1%
HLA-A*68:1 1%
HLA-A*68:2 1%
HLA-A*01:1
HLA-A*02:1
HLA-A*02:2
HLA-A*02:48
HLA-A*02:66
HLA-A*03:1
HLA-A*03:26
HLA-A*11:1
HLA-A*11:88
HLA-A*23:1
HLA-A*24:144
HLA-A*24:2
HLA-A*24:50
HLA-A*26:1
HLA-A*29:2
HLA-A*29:46
HLA-A*30:1
HLA-A*30:17
HLA-A*31:1
HLA-A*32:1
HLA-A*32:13
HLA-A*66:1
HLA-A*68:1
HLA-A*68:2
HLA-A*69:1


C
HLA-B Patients
HLA-B*07:02 11%
HLA-B*08:01 5%
HLA-B*13:02 3%
HLA-B*15:01 3%
HLA-B*15:17 3%
HLA-B*18:01 11%
HLA-B*18:26 5%
HLA-B*27:02 3%
HLA-B*27:05 5%
HLA-B*35:03 11%
HLA-B*35:24 5%
HLA-B*40:01 3%
HLA-B*40:02 3%
HLA-B*40:56 3%
HLA-B*44:02 11%
HLA-B*44:03 3%
HLA-B*50:01 3%
HLA-B*51:01 5%
HLA-B*52:01 3%
HLA-B*57:01 5%
HLA-B*07:02
HLA-B*08:01
HLA-B*13:02
HLA-B*15:01
HLA-B*15:17
HLA-B*18:01
HLA-B*18:26
HLA-B*27:02
HLA-B*27:05
HLA-B*35:03
HLA-B*35:24
HLA-B*40:01
HLA-B*40:02
HLA-B*40:56
HLA-B*44:02
HLA-B*44:03
HLA-B*50:01
HLA-B*51:01
HLA-B*52:01
HLA-B*57:01


D
HLA-B Control
HLA-B*07:2 7%
HLA-B*08:1 9%
HLA-B*08:109 1%
HLA-B*13:2 1%
HLA-B*14:1 2%
HLA-B*14:1 2%
HLA-B*14:2 1%
HLA-B*15:1 4%
HLA-B*15:17 2%
HLA-B*18:1 2%
HLA-B*18:30 2%
HLA-B*27:2 1%
HLA-B*27:5 6%
HLA-B*27:7 1%
HLA-B*35:1 6%
HLA-B*35:2 2%
HLA-B*35:8 1%
HLA-B*37:1 1%
HLA-B*38:1 2%
HLA-B*39:1 1%
HLA-B*39:6 1%
HLA-B*40:1 9%
HLA-B*40:2 3%
HLA-B*40:32 1%
HLA-B*41:1 1%
HLA-B*44:2 11%
HLA-B*44:3 3%
HLA-B*45:1 1%
HLA-B*49:1 2%
HLA-B*51:1 9%
HLA-B*52:1 2%
HLA-B*53:1 1%
HLA-B*55:1 1%
HLA-B*58:1 1%
HLA-B*58:2 1%
HLA-B*81:1 1%
HLA-B*07:2
HLA-B*08:1
HLA-B*08:109
HLA-B*13:2
HLA-B*14:1
HLA-B*14:1
HLA-B*14:2
HLA-B*15:1
HLA-B*15:17
HLA-B*18:1
HLA-B*18:30
HLA-B*27:2
HLA-B*27:5
HLA-B*27:7
HLA-B*35:1
HLA-B*35:2
HLA-B*35:8
HLA-B*37:1
HLA-B*38:1
HLA-B*39:1
HLA-B*39:6
HLA-B*40:1
HLA-B*40:2
HLA-B*40:32
HLA-B*41:1
HLA-B*44:2
HLA-B*44:3
HLA-B*45:1
HLA-B*49:1
HLA-B*51:1
HLA-B*52:1
HLA-B*53:1
HLA-B*55:1
HLA-B*58:1
HLA-B*58:2
HLA-B*81:1

E


HLA-C Patients
HLA-C*01:23 2%
HLA-C*02:02 13%
HLA-C*02:07 5%
HLA-C*03:03 3%
HLA-C*03:04 3%
HLA-C*04:01 8%
HLA-C*05:01 10%
HLA-C*06:021 16%
HLA-C*07:01 5%
HLA-C*07:02 8%
HLA-C*07:04 5%
HLA-C*12:02 3%
HLA-C*12:03 13%
HLA-C*14:02 3%
HLA-C*16:01 3%
HLA-C*01:23
HLA-C*02:02
HLA-C*02:07
HLA-C*03:03
HLA-C*03:04
HLA-C*04:01
HLA-C*05:01
HLA-C*06:021
HLA-C*07:01
HLA-C*07:02
HLA-C*07:04
HLA-C*12:02
HLA-C*12:03
HLA-C*14:02
HLA-C*16:01

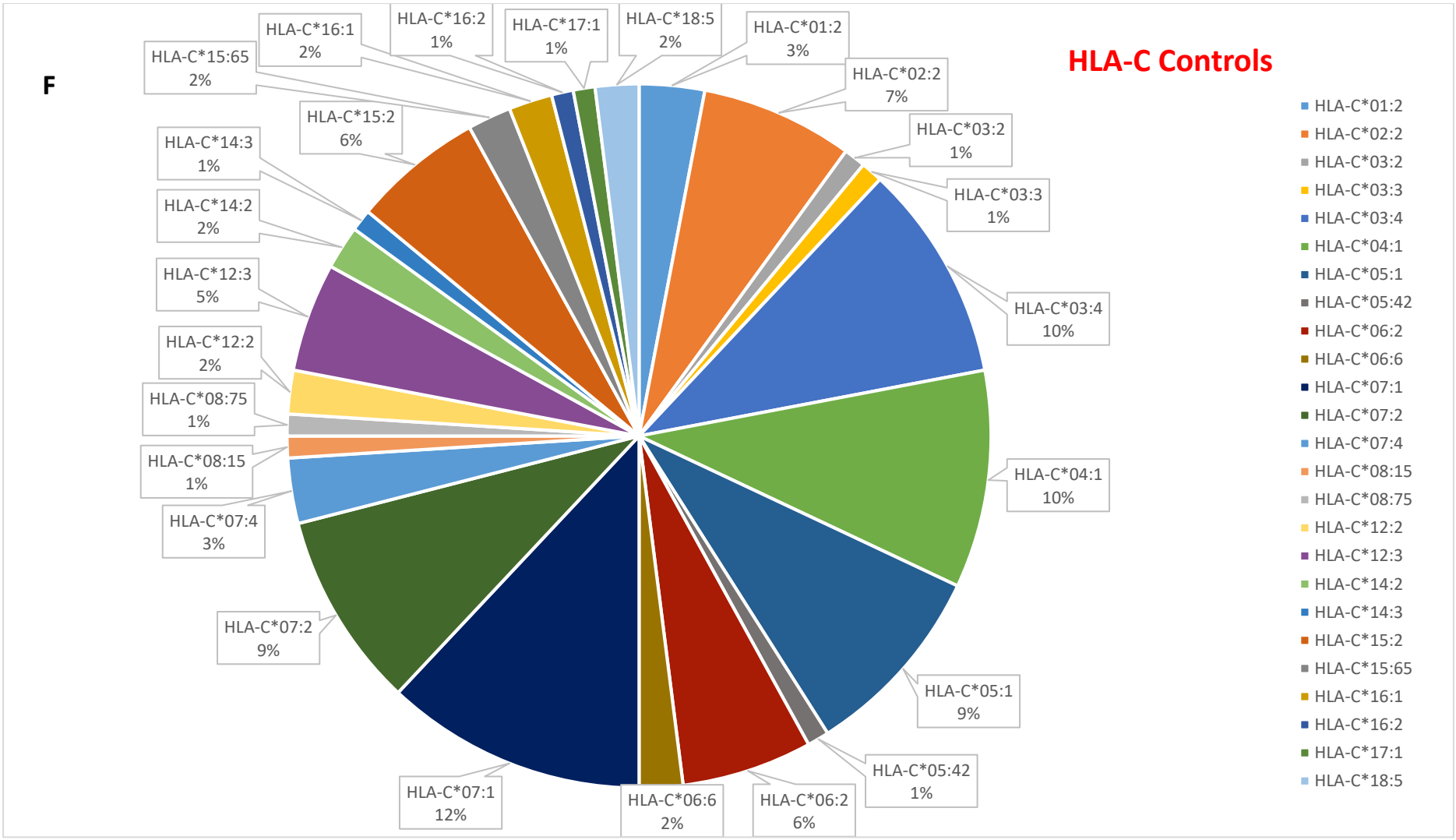


Figure 22. Distribution of allotypes A) HLA-A of cohort, B)HLA-A of 60 in-lab Swiss controls, C) HLA-B of cohort, D) HLA-B of Swiss controls E) HLA-C of cohort and F) HLA-C of Swiss controls

9. Related phenotypes

Skin presentation in SSc, and morphea including the LLS subtype is often similar, therefore a possible genetic relationship between the 2 diseases was investigated here (Ramos, Silver et al. 2015). SSc is not considered genetic in all patients, however several studies have identified some genetic factors involved such as certain genes and HLA allotypes. A list of 54 genes known to be involved in SSc was compiled from literature and then an analysis of how many of these genes were somatically affected in our patients was performed. *KCNA5*, a gene associated with pulmonary arterial hypertension in SSc patients, has been somatically affected with different mutations in 3 of 19 patients. These mutations were then found in blood in a lower allelic fraction of the patients indicating they are not truly somatic mutations.

10. Loss of heterozygosity

Output of LOH was analysed for a unifying event from all patient samples. Furthermore, an event of varying length, which was affecting the same area, was searched for. No LOH event was found to be present in more than one patient. LOH was ruled out of having possible involvement with LLS.

11. Somatic callers and deep sequencing

The analysis of the truncated bam files with inserted mutations (DP 20 and DP 100) showed VarScan to be a more restrictive calling program with the most lenient of filtering settings (VarScan Low Table 6) resulting in mutations at an allelic fraction below 15% to be filtered out at both DP. MuTect2 was found to be very sensitive as it is able to detect somatic mutations at an allelic fraction of 5% at both DP. Strelka was shown to be more sensitive at DP 100 where it could find somatic mutations at an allelic fraction between 5-10% as opposed to DP 20 where the lowest allelic fraction was 20% (Table 6).

Genotype caller	Allelic fraction											
	0.05	0.1	0.15	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
VarScan (High)	-	-	-	-	-	+	+	+	+	+	+	+
VarScan (Mid)	-	-	+	+	+	+	+	+	+	+	+	+
VarScan (Low)	-	-	+	+	+	+	+	+	+	+	+	+
Strelka (DP 20)	-	-	-	+	+	+	+	+	+	+	+	+
Strelka (DP 100)	-	+	+	+	+	+	+	+	+	+	+	+
Strelka (DP 100, Q5)	+	+	+	+	+	+	+	+	+	+	+	+
MuTect (DP 20)	+	+	+	+	+	+	+	+	+	+	+	+
MuTect (DP 100)	+	+	+	+	+	+	+	+	+	+	+	+

Table 6. Allelic fraction at which a mutation can be called as somatic by a calling program. VarScan investigated with DP20 sample with stringency setting Low to high. Strelka and MuTect investigated with both DP.

Number of mutations called	
Normal	**Deep sequencing**
116‘757	134‘773

Table 7. Number of mutations (SNPs and indels) called by haplotype caller in normal and deep sequenced samples

Following this, 4 patients were chosen for deep sequencing of both the blood and corresponding skin. Of these 4 patients, 2 had the skin biopsy taken from the *en coup de sabre* affected skin of the forehead. The other 2 patients had biopsies taken from the frontoparietal-temporal lesions. Deep sequencing samples were prepared for sequencing as before. Maximum sequencing depth was ensured by combining the raw data with the raw data from the previously sequenced samples of the same patients. A sequencing depth of 352.55 (± 26.8) for blood samples and 368.4 (± 67.7) for skin samples was achieved. Specificity of 85.15% in skin and 83.65% in blood was achieved. Somatic analysis and subsequent annotation of mutations were performed as above. Number of mutations called increased by 13.5% with the increased sensitivity (Table 7).

Despite the increased number of mutations, no rare and somatic SNP mutation was found in more than one patient and similarly no somatic indel was found in more than one patient. A gene based analysis was also performed to find genes affected in each patient with various rare and damaging mutations and indels. No gene was somatically affected in more than one patient with both a SNP and indel. One gene, *GOLGA6LA*, was affected with an indel in one sample and a SNP in another two samples, however further scrutiny into these mutations ruled them out as true somatic mutations and thus as a candidate gene.

Discussion

Patient data

Our clinical cohort consisted of 19 patients. 2 patients had no clinical data available aside from confirmation of LLS diagnosis and gender. Our study included 13 females and 6 males, the ratio was 2.2 to 1 (Table 8). This reflects the expected gender ratio that ranges from 2-6 to 1 in literature. Of the 17 patients, 2 patients had joint involvement. Both patients had lesions on the right limbs, one on the lower leg and the other on the upper arm. One patient had ocular involvement. This patient had ECDS on the right side affecting the face and extending into the scalp. Two other patients with ECDS had involvement of the CNS. Both these patients had lesions on the face and one had the lesion extending into the scalp (Table 12). This gives a total of 5 patients in the cohort with extracutaneous involvement. Studies such as Christen-Zaech 2008 show over 50% of LLS patients have some form of extracutaneous involvement in the CNS, musculoskeletal or gastrointestinal.

Unilateral lesions presented on the left in 6 patients and the right of 11 patients (Table 10). In literature there has been no evidence that one axis is more affected than the other. There were no cases of bilateral involvement in our cohort. This is to be expected, as it is quite rare. In a studies by (Zulian, Athreya et al. 2005, Christen-Zaech, Hakim et al. 2008), it was estimated that 11% of LLS patients have bilateral involvement. In these same studies ~50% of LLS patients had ECDS with the remaining having them on trunk and limbs. In our cohort, 6 patients had lesions on the trunk or limbs and 11 patients presented with ECDS. The largest proportion of patients had 2 lesions (n=8) though 5 lesions were found in 2 patients and 8 lesions were found in 1 (Table 10, 11).

Of the 17 patients, 7 did not receive systemic treatment. 10 received systemic treatment, 9 were treated with MTX and one with another, unspecified treatment. In the cohort, only 3 patients received topical medicine. The treatment received was a protopic ointment. All patients treated with protopic ointment also received MTX treatment. LLS treatment often requires aggressive systemic treatment and topical solutions. The 7 patients who did not receive systemic treatment had no other treatments (Table 13). Of these 7 patients, 3 had ECDS and the rest has trunk/limb lesions.

Average age of onset in the clinical cohort was 8.1 years (Table 9). This is slightly higher than the consensus of literature of 6 years. The delay in diagnosis was 1.4 years, which is reflective of a study by Weibel et al. 2011 which found the delay in diagnosis averages 1.5 years and can last up to 2 years before correctly diagnosed (Table 9) (Weibel, Laguda et al. 2011). Average LoSSI score of the patients included in this study was 5.4 with a range of 2 to 10 (Table 8). Few studies include data scoring disease severity or damage in each patient (LoSCAT, LoSSI, LoSDI, PGA-A/D) so a comparison of our patients to those in other cohorts is not possible.

Exonic analysis of patients

Specificity of capture from WES of both SF and FFPE as well as coverage was comparable in both sample types. In previous studies, FFPE samples are shown to have up to ~10-fold higher cytosine (C) to thymine (T) and this leads to sequencing artefacts. Sample treatment with UDG prior to DNA extraction has been shown to reverse deamination-related sequencing errors significantly (Do and Dobrovic 2012, Chen, Mosier et al. 2014). Inclusion of these two sample types to be the somatic tissue of patients was therefore approved.

It was shown in this study that the two sample types did not lead to any problems in sample handling, sequencing and subsequent analysis. The conclusion is that future studies should include the two sample types together after processing as described above. This will be highly beneficial as it can increase sample size and thus the power of analysis. One limitation is that the analysis was performed on samples that were in formalin for a maximum of 6 months. The effects of formalin over a longer period of time have not yet been shown.

Opposing the hypothesis put forward at the commencement of this study that LLS is caused by a somatic mutation following Blaschko's lines, no somatic mutation, neither SNP nor indel was found in common between the patients. This result is surprising. The rarity of the disease would be compatible with being caused by a shared single mutation. The adherence to Blaschko's lines would also suggest it is driven by a somatic mutation, which we did not find.

Furthermore a gene-based analysis also showed there was no single gene affected with a somatic mutation in common between the patients. The presence of various mutations in a common gene could also have accounted for the rarity and skin pattern however this was not the case. It was then hypothesized that the cause of LLS was not a single gene, but a pathway was disrupted in each patient. Each affected gene could be different but they share a common

pathway. This pathway-based analysis did not reveal any pathway to be disrupted in each patient that could be driving LLS.

Due to the lack of inheritance and adherence to Blaschko's lines it is not likely that the driving cause of LLS is a germline mutation. This being said, it was important to investigate this as a possible cause to perform a well-rounded study. However, this analysis was also important as it could have unearthed a possible genotype that increases susceptibility to LLS. There could be several factors, immunologic or metabolic, that could lead to a patient being more susceptible to LLS.

The analysis of the germline did not reveal any mutations that are not present in the controls that held up to statistical scrutiny and IGV visualization. All mutations with an observed p-value below but in close proximity to the expected p-value, were investigated. All mutations were found to be in controls in a very low allelic fraction. This shows that they are false positives that possibly arose from discrepancy in DP between cases and controls. Furthermore, in the gene-based analysis no gene was found to be affected with mutations that were not also present in controls. This indicates that no germline mutations in the exonic regions are driving LLS.

A karyotype refers to the size, shape and number of chromosomes in an individual. Altered karyotypic states have been implicated in many genetic diseases such as trisomies and persistent polyclonal B-cell lymphocytosis (PPBL) (Mossafa and Flandrin 2002). It was a possibility that LLS could be caused by a rare karyotypic event occurring in a mosaic pattern. A subgroup of the patients were chosen for CGH analysis. CGH is a tool used to identify duplications and deletions on a very large scale. All patients were shown to have healthy karyotypes indicating neither a mosaic or germline change in karyotype causes LLS.

Copy number variation results from sections of the genome being repeated and it varies from person to person due to duplications and deletions. It is a type of structural variation. It has been shown in literature to be inherited but somatic changes in CNV has been implicated in the development of several forms of cancer. Several copy number events were found in each patient but no gene or area was affected in all the patients or no event was in common in 2 or more patients. CNV events are not driving LLS in either a germline or somatic pattern (Favero, Joshi et al. 2015) (Thapar and Cooper 2013).

HLA is a gene complex that codes for the major histocompatibility complex (MHC) proteins. It is separated into three class types, MHC I, II and III. Class I and II control the immune system. Class III is involved in the complement system. They are arranged on cell surfaces and interact with immune cells. The MHC region on chromosome 6, is a highly polymorphic region resulting in many different alleles called allotypes. Different classes control interactions with different aspects of the immune system. MHC class I code for HLA-A, B, and C. MHC class II code for the most polymorphic regions HLA-DP, DM, DR, and DQ.

As part of this study, we wanted to investigate if LLS patients have a unifying allotype or a change in allotype between affected skin and germline. Our patients had no allotypic changes in skin that were not in blood. Furthermore no allotype appears to be associated with LLS as all the patients had various allotypes that is consistent with the distributions of allotypes in Switzerland (Buhler, Nunes et al. 2012). Allotypes that are known to be associated with SSc were investigated and no allotype was found in common (Ramos, Silver et al. 2015). A genetic comparison of SSc to LLS was also performed.

WES is a technique that has proved highly valuable in studies of SSc. Though it is still unknown what is the cause of SSc it has helped identify risk factors for SSc development e.g. 64 variants

were found in a phospholipid transporter gene, *ATP8B4* (Gao, Emond et al. 2016). In this study it was hypothesized that a connection could be made in the genetics between these two diseases because of the similar skin involvement and sclerosis that is indicative of both. No gene that has been implicated in SSc was found to be burdened with mutations in LLS. If there is a factor that connects these two diseases it does not lie in the exonic regions.

Sensitivity of the 3 somatic callers was analysed due to concerns of the small biopsy size (2-4 mm) as per guidelines of the study. This small size leads to difficulty separating skin layers (i.e. fibroblasts and keratinocytes). Blaschko's lines has been shown to be caused by the epidermal patterns of migration of cells (Moss, Larkins et al. 1993, Weibel and Harper 2008). After the biopsy is taken, a large portion is taken for histological confirmation, a crucial part of correct diagnosis. Left for DNA extraction was 1-2mm of skin. Separating layers of such a small sample was difficult and isolation of keratinocytes for cell culture was unsuccessful. Fibroblasts could be isolated and cultured, however fibroblasts are dermal cells. Deep sequencing has proven a useful method in diagnosis of neurological disorders through patient blood. In this study patients had a somatic mutation in the brain. Deep sequencing is used to increase the number of reads with the mutation and mutation detection programs were then able to pick up the mutation (Jamuar, Lam et al. 2014)

In this study, input DNA for WES has an unknown ratio of keratinocyte DNA to fibroblast DNA and other epidermal and dermal cells. Therefore, sensitivity of the somatic callers was determined to find the lowest allelic fraction at which a heterozygous mutation in a heterogeneous sample can be found. MuTect was especially sensitive when using both types of data. The sensitivity of the Strelka program benefited from using deep sequencing data because the sensitivity of detection increased from 20% at DP 20 to 5% at DP 100. A result of

5% at DP 100 means the programs can call mutations at a 5% allelic fraction. This means a cell type, representing 10% of the sample can have a heterozygous mutation that can be found by the callers. Samples that had been deep sequenced also had many more mutations called (>13%) when compared to sequencing data with normal coverage. This can therefore be used to compensate for sample heterogeneity.

After deep sequencing, several rare and damaging somatic SNPs were found in common between 2 of the 4 patients. Visualization using IGV showed that these mutations were present in the germline sample but at a lower DP possibly causing the level of false positivity. Some mutations were also shown to be present in uninvolved control samples.

Taken together, no somatic or germline SNP or indel is driving LLS and no gene is affected in all patients. Furthermore, no karyotypic event, copy number change or HLA allotype has been shown to be involved in the disease in any capacity.

Transcriptomic analysis performed by a collaborator on 4 patients did not reveal any protein to have increased or decreased expression in LLS skin when compared to blood. This analysis is indicative that the disease-causing factor is not at the mRNA level. However, this study is of very few patients so this experiment should be repeated by future studies.

Possible future studies should focus on changes in epigenetics. As yet there are no examples of diseases following BL that are caused by epigenetic factors. However, epigenetics is quite a new field, the importance of which has just started to be unearthed. Future studies should also focus on the intronic regions. Whole genome sequencing allows for the analysis of exonic and intronic data. The introns are estimated to only have 15% of disease-causing mutations but these areas have numerous repeats and are quite large resulting in a very difficult analysis (Bao, Huang et al. 2014). However, within the intronic regions lie regulatory sites, binding

sites and splice sites amongst other. The role of the intron is often forgotten as most mutations are found in exonic regions but mutations here can cause several downstream maleffects.

The choice to study LLS as opposed to other forms of morphea lies in the fact that it is the only subtype that follows BL. As other diseases following BL have been shown to be caused by somatic mosaicism the initial hypothesis of this study was that LLS would also exhibit this genetic architecture. Studying somatic mosaicism is preferable to simple germline analysis. Performing a germline analysis on a rare disease is difficult as one needs a large cohort and a large number of controls so the data holds up to statistical scrutiny. The benefit of a somatic analysis is the internal controls. Patients are not only compared to uninvolved exomes or genomes but there is further comparison of the affected tissue of the patient to the unaffected. The use of internal and external controls makes mosaic diseases preferable to study.

Pertaining to the field of morphea and LLS research, this study is very important. Though no genetic factor has been pinpointed as the cause of LLS, several causative factors have been ruled out. The study focused solely on LLS, however it is a subtype of morphea. LLS is linked to the other types of morphea in many ways often making them difficult to distinguish. A genetic analysis of one subtype is therefore informative about morphea as a whole. This work is the first genetic analysis of a subtype of morphea, so it will hopefully benefit and provide guidance to future researchers.

Code	DOB	Gender	FFPE	SF	LoSSI Score	Clinical data available
[01002]	04.12.2005	m		X	7	1
[01003]	22.07.2003	f		X	4	1
[01004]	15.06.2001	f		X	NA	1
[01005]	23.11.2005	f	X		10	1
[01006]	06.11.2000	m	X		3	1
[01008]	02.06.2008	m	X		2	1
[04010]	29.01.1974	m		X	NA	0
[02011]	04.06.1998	m	X		NA	1
[03012]	19.08.1989	f		X	8	1
[01013]	13.07.2003	f	X		2	1
[01014]	21.10.2008	f	X		3	1
[01015]	07.03.1989	f	X		NA	1
[02017]	27.09.2004	f		X	8	1
[04018]	15.07.2002	f	X		6	1
[01019]	04.06.2004	f		X	4	1
[01020]	16.07.1997	f	X		NA	0
[02021]	30.10.1987	f		X	8	1
[01023]	11.09.2003	f		X	5	1
[01024]	01.01.2012	m		X	5	1

Table 8. Clinical data of LLS study cohort. 1 = yes, 0 = no

Code	Age at inclusion (months)	Age at inclusion (years)	Date of 1st manifestation	Age at 1st manifestation (months)	Age at 1st manifestation (years)	Date of 1st diagnosis	Diagnostic delay (months)	Diagnostic delay (years)
[01002]	106	8.71	01.06.2011	66.83	5.49	15.03.2013	21.77	1.79
[01003]	134.87	11.08	01.01.2009	66.33	5.45	03.12.2013	59.9	4.92
[01004]	161.37	13.26	01.01.2004	31	2.55	01.01.2005	12.2	1
[01005]	109.5	9	01.07.2013	92.57	7.61	03.02.2014	7.23	0.59
[01006]	170.4	14.01	01.06.2007	79.93	6.57	01.06.2007	0	0
[01008]	78.5	6.45	01.06.2010	24.3	2	20.05.2011	11.77	0.97
[04010]	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
[02011]	199.97	16.44	01.01.2000	19.2	1.58	01.01.2000	0	0
[03012]	307.13	25.24	01.01.2002	150.6	12.38	01.01.2006	48.7	4
[01013]	138	11.34	01.01.2008	54.43	4.47	01.01.2010	24.37	2
[01014]	74.07	6.09	01.04.2011	29.73	2.44	03.02.2012	10.27	0.84
[01015]	315.07	25.9	01.08.2014	309.27	25.42	05.12.2014	4.2	0.35
[02017]	126	10.36	01.01.2008	39.7	3.26	01.01.2012	48.7	4
[04018]	154.1	12.67	01.03.2014	141.57	11.64	05.11.2014	8.3	0.68
[01019]	133.23	10.95	01.09.2014	124.7	10.25	23.04.2015	7.8	0.64
[01020]	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
[02021]	338.7	27.84	01.01.2011	282.13	23.19	01.01.2011	0	0
[01023]	151.63	12.46	01.12.2014	136.63	11.23	01.12.2015	12.17	1
[01024]	52.13	4.28	01.01.2015	36.53	3	08.04.2016	15.43	1.27
	mean= 161.8	mean = 13.3		mean = 99.1	mean = 8.1		mean = 17.2	mean = 1.4
	median = 138	median = 11		median = 67	median = 5		median = 12	median = 1

Table 9. Clinical data pertaining to manifestation and delay in diagnosis of LLS study cohort

Code	Type of morphea	Right/left	Body region 1	Body region 2
[01002]	1	1	1	
[01003]	1	2	1	
[01004]	0	1	3	
[01005]	0	1	2	3
[01006]	1	1	1	
[01008]	1	2	1	
[04010]				
[02011]	1	1	1	
[03012]	0	1	2	3
[01013]	1	2	1	
[01014]	1	2	1	
[01015]	0	1	3	
[02017]	1	1	1	
[04018]	0	1	3	
[01019]	1	1	1	
[01020]				
[02021]	0	2	2	3
[01023]	1	1	1	
[01024]	1	2	1	

Table 10. Affected axis of patients in LLS study cohort. Type of morphea: 0 = linear limb/trunk, 1 = ECDS. Right/left: 1 = right, 2 = left. Body region: 1 = head, 2 = trunk, 3 = extremities.

Code	Location 1	Location 2	Location 3	Location 4	Location 5	Location 6	Location 7	Location 8	Number of lesions
[01002]	1	2							2
[01003]	1	2							2
[01004]	7	8	9						3
[01005]	3	5	7	8	9	10	11	12	8
[01006]	1	2							2
[01008]	2								1
[04010]									
[02011]	1	2							2
[03012]	3	7	8	9	10				5
[01013]	2								1
[01014]	2								1
[01015]	11								1
[02017]	1	2							2
[04018]	10	11							2
[01019]	1	2							2
[01020]									
[02021]	4	5	7	8	9				5
[01023]	1	2							2
[01024]	2								1

Table 11. Location and number of lesions of LLS study cohort. 1 = scalp, 2 = face, 3 = chest, 4 = abdomen, 5 = upper back, 6 = lower back, 7 = upper arm, 8 = forearm, 9 = hand, 10 = thigh, 11 = lower leg, 12 = foot

Code	Extracutaneous manifestations			
	Joints	**Eyes**	**CNS**	**Other autoimmune disease**
[01002]	0	1	0	0
[01003]	0	0	1	0
[01004]	1	0	0	0
[01005]	0	0	0	1
[01006]	0	0	0	0
[01008]	0	0	1	0
[04010]				
[02011]	0	0	0	0
[03012]	0	0	0	0
[01013]	0	0	0	0
[01014]	0	0	0	0
[01015]	1	0	0	0
[02017]	0	0	0	1
[04018]	0	0	0	0
[01019]	0	0	0	0
[01020]				
[02021]	0	0	0	0
[01023]	0	0	0	0
[01024]	0	0	0	0

Table 12. Extracutaneous manifestations of LLS study cohort. 1 = yes and 0 = no

Code	Current treatment details		
	Tx at inclusion	**Systemic medication**	**Topical medication**
[01002]	3	1	2
[01003]	1	1	0
[01004]	1	3	0
[01005]	3	1	2
[01006]	0	0	0
[01008]	0	0	0
[04010]			
[02011]	1	1	0
[03012]	0	0	0
[01013]	0	0	0
[01014]	3	1	2
[01015]	0	0	0
[02017]	1	1	0
[04018]	1	1	0
[01019]	0	0	0
[01020]			
[02021]	1	1	0
[01023]	0	0	0
[01024]	1	1	0

Table 13. Available data on treatment of LLS study cohort. Treatment at inclusion: 0 = no, 1 = systemic, 2 = topical, 3 = systemic and topical. Systemic medications: 0 = no systemic medication, 1 = MTX, 2 = MTX + systemic steroids, 3 = other. Topical medications: 0 = no topical med., 1 = topical steroids, 2 = protopic ointment

Chapter 2

Generalized Comedones, Acne, and Hidradenitis Suppurativa in a Patient with a *FGFR2* Missense mutation

Abbreviations

CD	Cluster of differentiation
CADD	Combined Annotation Dependent Depletion
HS	Hidradenitis suppurativa
IGV	Integrative genome viewer
MAF	Minor allele frequency
NC	Nevus comedonicus
NCS	Nevus comedonicus syndrome
PUVA	Psoralen and ultraviolet A

Abstract

Mutations in the fibroblast growth factor-receptor gene 2 (*FGFR2*) have been implicated in numerous diseases, including nevus comedonicus (NC) and naevoid acne that have somatic missense mutations in *FGFR2* in the affected tissue. A patient presented in our department with unusual, innumerable and large comedones throughout his back reminiscent of NC, as well as multifocal hidradenitis suppurativa and acne. Topical and systemic treatments were unsuccessful. Whole exome sequencing of blood-derived DNA detected a germline mutation in *FGFR2* that was predicted to be damaging. This could explain the multifocal and severe nature of the disease. We suggest screening other, phenotypically similar patients, for *FGFR2* mutations. Our findings, once confirmed independently, could indicate that therapeutic modulation of FGFR signalling in the acne tetrad could be effective.

Introduction

Hidradenitis suppurativa (HS) is an inflammatory skin disease that affects roughly 1% of Europeans. It is characterized by painful abscesses, boils, cysts, and malodorous, pus-filled lesions in intertriginous regions (Pink, Simpson et al. 2012, Pink, Simpson et al. 2013). It can follow autosomal dominant inheritance in some families, and mutations have been reported (Wang, Yang et al. 2010) in the gamma-secretase genes Presenilin-1 (*PSEN1*), Presenilin Enhancer-2 (*PSENEN*), and Nicastrin (*NCSTN*) in a minority (<7%) of cases. The remaining cohort are presumed to be sporadic or driven by other, as yet unidentified mutations.

Nevus comedonicus (NC) is a rare (prevalence 1:45,000–1:100,000) epidermal nevus comprising of a group of dilated hair follicle openings filled with plugs of brownish-black oxidized keratin (Kofmann 1895). It is usually localized on the head and neck and was initially described as “localized acne” (Munro and Wilkie 1998). If it occurs as a part of the nevus comedonicus syndrome (NCS) (Engber 1978), it can be generalized and form linear streaks. HS-like lesions have been observed in NCS (Qian, Liu et al. 2015). In some cases of NC, a somatic mutation in fibroblast growth factor-receptor gene 2 (*FGFR2*) has been identified, namely the Ser252Trp missense mutation (Munro and Wilkie 1998). *FGFR2* is expressed in keratinocytes, hair follicles, and sebaceous glands and has been implicated to induce hypercornification and comedogenesis (Melnik 2009). Germline *FGFR2* mutations are also associated with acne, as seen in dominant Apert syndrome (Campanati, Marconi et al. 2002) (also comprising craniosynostosis, epiphyseal closure, and syndactyly) (Moloney, Slaney et al. 1996, Downs, Condon et al. 1999). When occurring as a mosaic, the acne lesions follow Blaschko’s lines (Kiritsi, Lorente et al. 2015), the pattern of embryological cell development

and proliferation. In a recent case, a postzygotic mosaicism was found in exon 4 of *FGFR2* (c.758C>G, p.Pro253Arg) in low copy number. A similar mosaicism was also found in two other patients in p.Ser252Trp of exon 4 (Melnik 2009).

A 49-year-old male construction worker presented with a 29-year history of multiple abscesses in his buttock, inguinal, and axillary regions (Figure 23) and a history of severe post-pubertal acne conglobate. His back was studded with innumerable large comedones not unlike those seen in NC, however rather more spread out (Figure 23A), and with significant scarring. His condition had proved resistant to treatment with topical and systemic antibiotics, dapsone, retinoids, including acitretin and isotretinoin, zinc supplements and PUVA. Treatment with infliximab did have an effect on the inflammatory components; however, it was minimal and ultimately unsuccessful. Surgical interventions included excision of the left axilla and shoulder region and a transposition flap and excision of a pilonidal sinus.

The patient is an orphan. It is known that several other people in his family including his father and brother had suffered from the same condition with different intensities; however, there is no contact between the patient and family. Requests to collect samples from family members were denied by the patient. Suspecting a potential genetic origin (Fitzsimmons, Guilbert et al. 1985), we therefore performed whole exome sequencing on blood-derived DNA.

Figure 23. Widespread abscesses present in 49-year-old male patient.
A) Back of patient showing grouped comedones and hidradenitis suppurativa (HS) scars.
B) Inguinal scars.
C) Close-up of open HS lesions.

Materials and Methods

DNA was extracted from blood peripheral blood mononuclear cells as previously described in the “Processing of biological samples” section of materials and methods. Whole exome sequencing was performed as described in previous section “Whole exome sequencing” and bioinformatic analysis and quality control were performed as per directions in the “Bioinformatics analysis” subsections Genome analysis toolkit section and annotation and quality control of materials and methods. Bioinformatic analysis was performed with the aim of analysing the patient as an individual or “Single variant” analysis (Figure 17).

All rare (MAF < 0.01) and damaging mutations (CADD >15) were extracted for analysis. First genes that are known to be associated with acne were analysed for mutations. All variants in these genes were not damaging or rare. All rare and damaging mutations were visualized, quality controlled and compared to controls without acne using IGV. All mutations deemed truly rare were entered into the Online Mendelian Inheritance in Man (OMIM) database. Furthermore, a literary search was performed.

Results

A search for rare, deleterious mutations revealed that there was no such mutation present in any of the known HS-associated genes. However, a rare missense mutation in the *FGFR2* gene was found in exon 5 (c.G492C, p.K164N, see table 14). This mutation is predicted (McLaren, Pritchard et al. 2010) to have pathological consequence on the protein by several prediction algorithms, namely SIFT ("deleterious," lowest score 0), PolyPhen ("probably damaging," highest score 1), LoFtool ("probably damaging," 0.00179), Condel ("deleterious," 1.00), CADD (17.05), and SNPs&Go (73% probability). MutPred resulted in 63% probability for "damaging," and predicted (Li, Krishnan et al. 2009) molecular features of the mutation are loss of methylation at K164 (P = 0.0069), loss of ubiquitination at K164 (P = 0.0081), loss of solvent accessibility (P = 0.0371), and loss of MoRF binding (not significant, P = 0.0575) as well as gain of sheet (not significant, P = 0.1451).

Discussion

In a patient with clinical features reminiscent of NCS, we found a new, rare heterozygous missense mutation in *FGFR2*, the gene for the keratinocyte growth factor (KGF) receptor. *FGFR2*, also known as CD332, was cloned in 1990 (Houssaint, Blanquet et al. 1990) and soon found to be the receptor for KGF. It has an extracellular part constructed of three immunoglobulin domains, a lipophilic transmembrane part, and a tyrosine kinase that extends in the cytoplasm. Fibroblast growth factors (FGF) can bind the extracellular domain and thus activate the tyrosine kinase. This leads to cell division and differentiation. The isoform found in ectodermal tissues such as the skin is the FGFR2IIIb, also known as the KGF receptor. It binds FGF-7 and 10, also known as KGF 1 and 2, as well as FGF-1, -3, -22. Keratinocyte growth factors are potent mitogens for many epithelial cell types but lack detectable activity on fibroblasts or endothelial cells.

Germline mutations in *FGFR2* underlie various craniosynostosis syndromes (Table 15), and somatic mutations in the same gene have been reported in NC (Munro and Wilkie 1998) and nevoid acne (Campanati, Marconi et al. 2002). We report a germline mutation in *FGFR2* that may contribute to explaining a generalized, rather than mosaic, NC-like, acneiform phenotype with additional features of HS. As documented above, the disease proved highly treatment resistant, which is a common feature in NC. FGFR inhibitors, which may directly address the genetically driven pathogenesis, exist but have not been reported in this context.

Chr	Pos	Ref	Alt	Exon	Mutation	MAF in 1KG, EVS, ExAC
10	123310936	C	G	exon5	c.G492C, p.K164N	0.00
SIFT	**PolyPhen**	**Condel**	**LoFtool**	**SNPs&Go**	**MutPred**	**CADD**
deleterious	probably_damaging	deleterious (1.0)	probably damaging (0.00179)	Disease (0.73)	Disease (0.626)	17.1

Table 14. Fibroblast growth factor-receptor gene 2 (FGFR2) mutation identified in the patient.

Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key
10q26.13	Antley–Bixler syndrome without genital anomalies or disordered steroidogenesis	207410	AR	3
10q26.13	Apert syndrome	101200	AD	3
10q26.13	Beare-Stevenson cutis gyrata syndrome	123790	AD	3
10q26.13	Bent bone dysplasia syndrome	614592	AD	3
10q26.13	Craniofacial-skeletal-dermatologic dysplasia	101600	AD	3
10q26.13	Craniosynostosis, non-specific			3
10q26.13	Crouzon syndrome	123500	AD	3
10q26.13	Gastric cancer, somatic	613659		3
10q26.13	Jackson–Weiss syndrome	123150	AD	3
10q26.13	LADD syndrome	149730	AD	3
10q26.13	Pfeiffer syndrome	101600	AD	3
10q26.13	Saethre–Chotzen syndrome	101400	AD	3
10q26.13	Scaphocephaly and Axenfeld-Rieger anomaly		NA	3
10q26.13	Scaphocephaly, maxillary retrusion, and mental retardation	609579	NA	3
10q26.13	Nevus comedonicus	–	Somatic	none

Table 15. Diseases caused by *FGFR2* mutation

References

Abraham, D., et al. (2009). "Overview of pathogenesis of systemic sclerosis." Rheumatology **48**(suppl_3): iii3-iii7.

Amaral, T. N., et al. (2012). "Neurologic involvement in scleroderma en coup de sabre." Autoimmune diseases **2012**.

Amaral, T. N., et al. (2013). Neurologic involvement in scleroderma: a systematic review. Seminars in arthritis and rheumatism, Elsevier.

Ana-Maria, F., et al. (2008). "Disabling pansclerotic morphea of childhood–unusual case and management challenges." Journal of medicine and life **1**(3): 348.

Arkachaisri, T., et al. (2009). "The localized scleroderma skin severity index and physician global assessment of disease activity: a work in progress toward development of localized scleroderma outcome measures." The Journal of rheumatology **36**(12): 2819-2829.

Asadollahi, R., et al. (2014). "The clinical significance of small copy number variants in neurodevelopmental disorders." Journal of medical genetics: jmedgenet-2014-102588.

Auwera, G. A., et al. (2013). "From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline." Current Protocols in Bioinformatics: 11.10. 11-11.10. 33.

Bao, R., et al. (2014). "Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing." Cancer informatics **13**(Suppl 2): 67.

Beaudet, A. L. (2013). "The utility of chromosomal microarray analysis in developmental and behavioral pediatrics." Child development **84**(1): 121-132.

Bendeck, S. E. and H. T. Jacobe (2007). "Ultrasound as an outcome measure to assess disease activity in disorders of skin thickening: an example of the use of radiologic techniques to assess skin disease." Dermatologic therapy **20**(2): 86-92.

Biesecker, L. G. and N. B. Spinner (2013). "A genomic view of mosaicism and human disease." Nature Reviews Genetics **14**(5): 307-320.

Birdi, N., et al. (1992). "Childhood linear scleroderma: a possible role of thermography for evaluation." The Journal of rheumatology **19**(6): 968-973.

Blaschko (1901). Die Nervenverteilung in der Haut in ihrer Beziehung zu den Erkrankungen der Haut: Beilage zu den Verhandlungen der Deutschen Dermatologischen Gesellschaft, VII. Congress zu reslau im Mai 1901, Braumüller.

Blaszczyk, M., et al. (2003). "Progressive facial hemiatrophy: central nervous system involvement and relationship with scleroderma en coup de sabre." The Journal of rheumatology **30**(9): 1997-2004.

Buhler, S., et al. (2012). "The heterogeneous HLA genetic makeup of the Swiss population." PloS one **7**(7): e41400.

Burrows-Wheeler-Aligner-kit (2016).

Campanati, A., et al. (2002). "Pronounced and early acne in Apert's syndrome: a case successfully treated with oral isotretinoin." European journal of dermatology: EJD **12**(5): 496-498.

Careta, M. F. and R. Romiti (2015). "Localized scleroderma: clinical spectrum and therapeutic update." Anais brasileiros de dermatologia **90**(1): 62-73.

Chen, G., et al. (2014). "Cytosine deamination is a major cause of baseline noise in next-generation sequencing." Molecular diagnosis & therapy **18**(5): 587-593.

Chen, R., et al. (2015). "Whole-exome enrichment with the Agilent SureSelect human all exon platform." Cold Spring Harbor Protocols **2015**(7): pdb. prot083659.

Chiu, Y. E., et al. (2012). "A Significant Proportion of Children with Morphea En Coup De Sabre and Parry-Romberg Syndrome Have Neuroimaging Findings." Pediatric dermatology **29**(6): 738-748.

Christen-Zaech, S., et al. (2008). "Pediatric morphea (localized scleroderma): review of 136 patients." Journal of the American Academy of Dermatology **59**(3): 385-396.

Cibulskis, K., et al. (2013). "Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples." Nature biotechnology **31**(3): 213-219.

DePristo, M. A., et al. (2011). "A framework for variation discovery and genotyping using next-generation DNA sequencing data." Nature genetics **43**(5): 491-498.

Do, H. and A. Dobrovic (2012). "Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil-DNA glycosylase." Oncotarget **3**(5): 546-558.

Do, H. and A. Dobrovic (2012). "Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil-DNA glycosylase." Oncotarget **3**(5): 546.

Dobrev, H. (2005). "Application of Cutometer area parameters for the study of human skin fatigue." Skin Research and Technology **11**(2): 120-122.

Downs, A., et al. (1999). "Isotretinoin therapy for antibiotic-refractory acne in Apert's syndrome." Clin Exp Dermatol **24**(6): 461-463.

El-Kehdy, J., et al. (2012). "A review of Parry-Romberg syndrome." Journal of the American Academy of Dermatology **67**(4): 769-784.

Elhai, M., et al. (2011). "Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies." Rheumatology **51**(6): 1017-1026.

Engber, P. B. (1978). "The nevus comedonicus syndrome: a case report with emphasis on associated internal manifestations." International journal of dermatology **17**(9): 745-749.

Ewing, A. D., et al. (2015). "Combining tumor genome simulation with crowdsourcing to benchmark somatic single-nucleotide-variant detection." Nature methods **12**(7): 623-630.

Falanga, V., et al. (1986). "Linear SclerodermaClinical Spectrum, Prognosis, and Laboratory Abnormalities." Annals of internal medicine **104**(6): 849-857.

Favero, F., et al. (2015). "Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data." Annals of Oncology **26**(1): 64-70.

Fett, N. and V. P. Werth (2011). "Update on morphea: part I. Epidemiology, clinical presentation, and pathogenesis." Journal of the American Academy of Dermatology **64**(2): 217-228.

Fett, N. and V. P. Werth (2011). "Update on morphea: part II. Outcome measures and treatment." Journal of the American Academy of Dermatology **64**(2): 231-242.

Fett, N. M. (2013). "Morphea (localized scleroderma)." JAMA dermatology **149**(9): 1124-1124.

Fitzsimmons, J., et al. (1985). "Evidence of genetic factors in hidradenitis suppurativa." British Journal of Dermatology **113**(1): 1-8.

Fleischmajer, R., et al. (1977). "Cellular infiltrates in scleroderma skin." Arthritis & Rheumatology **20**(4): 975-984.

França, L. T., et al. (2002). "A review of DNA sequencing techniques." Quarterly reviews of biophysics **35**(2): 169-200.

Gambichler, T., et al. (2001). "Bilateral linear scleroderma" en coup de sabre" associated with facial atrophy and neurological complications." BMC dermatology **1**(1): 9.

Gambichler, T., et al. (2003). "Linear scleroderma 'en coup de sabre'treated with topical calcipotriol and cream psoralen plus ultraviolet A." Journal of the European Academy of Dermatology and Venereology **17**(5): 601-602.

Gao, L., et al. (2016). "Identification of Rare Variants in ATP8B4 as a Risk Factor for Systemic Sclerosis by Whole-Exome Sequencing." Arthritis & Rheumatology **68**(1): 191-200.

Garcia-Romero, M. T., et al. (2017). "The role of local temperature and other clinical characteristics of localized scleroderma as markers of disease activity." International journal of dermatology **56**(1): 63-67.

Go, J. L., et al. (2001). The trigeminal nerve. Seminars in Ultrasound, CT and MRI, Elsevier.

Gruss, C. J., et al. (2001). "Effects of low dose ultraviolet A-1 phototherapy on morphea." Photodermatology, photoimmunology & photomedicine **17**(4): 149-155.

Happle, R. and A. Assim (2001). "The lines of Blaschko on the head and neck." Journal of the American Academy of Dermatology **44**(4): 612-615.

Harboe, T. L., et al. (2011). "Mosaicism in segmental Darier disease: an in-depth molecular analysis quantifying proportions of mutated alleles in various tissues." Dermatology **222**(4): 292-296.

Hardy, J., et al. (2016). Pansclerotic morphea with a rapidly fatal outcome in an 11-year-old girl. Annales de dermatologie et de venereologie.

Henderson, C., et al. (1995). "Apert's syndrome and androgen receptor staining of the basal cells of sebaceous glands." British Journal of Dermatology **132**(1): 139-143.

Henoch, E. and M. H. Romberg (1846). Klinische Ergebnisse, A. Foerstner.

Highnam, G., et al. (2015). "An analytical framework for optimizing variant discovery from personal genomes." Nature communications **6**.

Hoesly, J. M., et al. (1987). "Localized scleroderma (morphea) and antibody to Borrelia burgdorferi." Journal of the American Academy of Dermatology **17**(3): 455-458.

Houssaint, E., et al. (1990). "Related fibroblast growth factor receptor genes exist in the human genome." Proceedings of the national academy of sciences **87**(20): 8180-8184.

Huh, W., et al. (2007). "Congenital acantholytic dyskeratotic epidermal naevus following Blaschko's lines versus segmental Darier's disease." European Journal of Dermatology **17**(2): 130-132.

Hykin, S. M., et al. (2015). "Fixing formalin: a method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing." PloS one **10**(10): e0141579.

Jamuar, S. S., et al. (2014). "Somatic mutations in cerebral cortical malformations." New England Journal of Medicine **371**(8): 733-743.

Johnson, R. and M. Ziff (1976). "Lymphokine stimulation of collagen accumulation." Journal of clinical Investigation **58**(1): 240.

Joly, P., et al. (1994). "Treatment of severe forms of localized scleroderma with oral corticosteroids: follow-up study on 17 patients." Archives of dermatology **130**(5): 663-664.

Jun, J. H., et al. (2011). "Parry-romberg syndrome with en coup de sabre." Annals of dermatology **23**(3): 342-347.

Kang, H. M. "EPACTS: efficient and parallelizable association container toolbox. 2014." Google Scholar.

Karaaltin, M. V., et al. (2012). "Treatment of “en coup de sabre” deformity with adipose-derived regenerative cell–enriched fat graft." Journal of Craniofacial Surgery **23**(2): e103-e105.

Kircher, M., et al. (2014). "A general framework for estimating the relative pathogenicity of human genetic variants." Nature genetics **46**(3): 310-315.

Kiritsi, D., et al. (2015). "Blaschko line acne on pre-existent hypomelanosis reflecting a mosaic FGFR2 mutation." British Journal of Dermatology **172**(4): 1125-1127.

Koboldt, D. C., et al. (2012). "VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing." Genome research **22**(3): 568-576.

Kofmann, S. (1895). "Ein fall von seltener localisation und verbreitung von comedonen." Archives of Dermatological Research **32**(1): 177-178.

Kreuter, A., et al. (2001). "Combined Treatment with Calcipotriol Ointment and Low-Dose Ultraviolet A1 Phototherapy in Childhood Morphea." Pediatric dermatology **18**(3): 241-245.

Kreuter, A., et al. (2009). "Diagnosis and therapy of localized scleroderma." JDDG: Journal der Deutschen Dermatologischen Gesellschaft **7**(s6).

Kreuter, A., et al. (2016). "German guidelines for the diagnosis and therapy of localized scleroderma." JDDG: Journal der Deutschen Dermatologischen Gesellschaft **14**(2): 199-216.

Laxer, R. M. and F. Zulian (2006). "Localized scleroderma." Current opinion in rheumatology **18**(6): 606-613.

Li, B., et al. (2009). "Automated inference of molecular mechanisms of disease from amino acid substitutions." Bioinformatics **25**(21): 2744-2750.

Li, H. and R. Durbin (2010). "Fast and accurate long-read alignment with Burrows–Wheeler transform." Bioinformatics **26**(5): 589-595.

Li, H., et al. (2009). "The sequence alignment/map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Li, S. C., et al. (2010). "Treatment of pediatric localized scleroderma: results of a survey of North American pediatric rheumatologists." The Journal of rheumatology **37**(1): 175-181.

Liu, B., et al. (2013). "Computational methods for detecting copy number variations in cancer genome using next generation sequencing: principles and challenges." Oncotarget **4**(11): 1868.

Majewski, J., et al. (2011). "What can exome sequencing do for you?" Journal of medical genetics **48**(9): 580-589.

Major, E., et al. (2013). "HLA typing from 1000 genomes whole genome and whole exome illumina data." PloS one **8**(11): e78410.

Malane, M. S., et al. (1991). "Diagnosis of Lyme disease based on dermatologic manifestations." Annals of internal medicine **114**(6): 490-498.

Mardis, E. R. (2008). "Next-generation DNA sequencing methods." Annu. Rev. Genomics Hum. Genet. **9**: 387-402.

Marsol, I. B. (2013). "Update on the classification and treatment of localized scleroderma." Actas Dermo-Sifiliográficas (English Edition) **104**(8): 654-666.

Martini, G., et al. (2009). "Successful treatment of severe or methotrexate-resistant juvenile localized scleroderma with mycophenolate mofetil." Rheumatology **48**(11): 1410-1413.

McKenna, A., et al. (2010). "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." Genome research **20**(9): 1297-1303.

McLaren, W., et al. (2010). "Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor." Bioinformatics **26**(16): 2069-2070.

Melnik, B. (2009). "Role of FGFR2-signaling in the pathogenesis of acne." Dermato-endocrinology **1**(3): 141-156.

Mertens, J. S., et al. (2016). "Use of mycophenolate mofetil in patients with severe localized scleroderma resistant or intolerant to methotrexate." Acta dermato-venereologica **96**(4): 510-513.

Metzker, M. L. (2010). "Sequencing technologies—the next generation." Nature Reviews Genetics **11**(1): 31.

Molho-Pessach, V. and J. V. Schaffer (2011). "Blaschko lines and other patterns of cutaneous mosaicism." Clinics in dermatology **29**(2): 205-225.

Moloney, D. M., et al. (1996). "Exclusive paternal origin of new mutations in Apert syndrome." Nature genetics **13**(1): 48.

Moss, C., et al. (1993). "Epidermal mosaicism and Blaschko's lines." Journal of medical genetics **30**(9): 752-755.

Mossafa, H. and G. Flandrin (2002). Persistent polyclonal B-cell lymphocytosis. Annales de biologie clinique.

Munro, C. S. and A. O. Wilkie (1998). "Epidermal mosaicism producing localised acne: somatic mutation in FGFR2." The Lancet **352**(9129): 704-705.

Nariai, N., et al. (2015). "HLA-VBSeq: accurate HLA typing at full resolution from whole-genome sequencing data." BMC genomics **16**(2): S7.

Neustadter, J. H., et al. (2009). "Extracorporeal photochemotherapy for generalized deep morphea." Archives of dermatology **145**(2): 127-130.

Nijhawan, R. I., et al. (2011). "Early localized morphea mimicking an acquired port-wine stain." Journal of the American Academy of Dermatology **64**(4): 779-782.

Orozco-Covarrubias, L., et al. (2002). "Scleroderma 'en coup de sabre'and progressive facial hemiatrophy. Is it possible to differentiate them?" Journal of the European Academy of Dermatology and Venereology **16**(4): 361-366.

Paller, A. S. (2004). "Piecing together the puzzle of cutaneous mosaicism." Journal of clinical Investigation **114**(10): 1407.

Peterson, L. S., et al. (1995). Classification of morphea (localized scleroderma). Mayo Clinic Proceedings, Elsevier.

Pink, A. E., et al. (2012). "Mutations in the γ-secretase genes NCSTN, PSENEN, and PSEN1 underlie rare forms of hidradenitis suppurativa (acne inversa)." The Journal of investigative dermatology **132**(10): 2459.

Pink, A. E., et al. (2013). "γ-Secretase mutations in hidradenitis suppurativa: new insights into disease pathogenesis." Journal of investigative dermatology **133**(3): 601-607.

Poblete-Gutiérrez, P., et al. (2004). "Allelic loss underlies type 2 segmental Hailey-Hailey disease, providing molecular confirmation of a novel genetic concept." Journal of clinical Investigation **114**(10): 1467.

Polcari, I., et al. (2014). "Headaches as a presenting symptom of linear morphea en coup de sabre." Pediatrics **134**(6): e1715-e1719.

Prinz, J. C., et al. (2009). "“Borrelia-associated early-onset morphea”: A particular type of scleroderma in childhood and adolescence with high titer antinuclear antibodies?: Results of a cohort analysis and presentation of three cases." Journal of the American Academy of Dermatology **60**(2): 248-255.

Qian, G., et al. (2015). "Naevus comedonicus syndrome complicated by hidradenitis suppurativa-like lesions responding to acitretin treatment." Acta dermato-venereologica **95**(8): 992-993.

Qin, M., et al. (2015). "SCNVSim: somatic copy number variation and structure variation simulator." BMC bioinformatics **16**(1): 66.

Rai, R., et al. (2000). "Bilateral en coup de sabre—a rare entity." Pediatric dermatology **17**(3): 222-224.

Ramos, P. S., et al. (2015). "Genetics of systemic sclerosis: recent advances." Current opinion in rheumatology **27**(6): 521-529.

Raposo-do-Amaral, C., et al. (2001). "Therapeutic approTherapeutic Approach To The Parry-Romberg Syndrome Based On A Severity Grading System." Revista Brasileira de Cirurgia Plástica **29**(1): 57-65.

Resende, L., et al. (1991). "Experimental study of progressive facial hemiatrophy: effects of cervical sympathectomy in animals." Revue neurologique **147**(8-9): 609-611.

Rodnan, G. P., et al. (1979). "Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma." Arthritis & Rheumatology **22**(2): 130-140.

Sakuntabhai, A., et al. (2000). "Mosaicism for ATP2A2 mutations causes segmental Darier's disease." Journal of investigative dermatology **115**(6): 1144-1147.

Sanger, F. and A. R. Coulson (1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase." Journal of molecular biology **94**(3): 441-448.

Sanger, F., et al. (1977). "DNA sequencing with chain-terminating inhibitors." Proceedings of the national academy of sciences **74**(12): 5463-5467.

Sartori-Valinotti, J. C., et al. (2013). "Updates on morphea: role of vascular injury and advances in treatment." Autoimmune diseases **2013**.

Saunders, C. J., et al. (2012). "Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units." Science translational medicine **4**(154): 154ra135-154ra135.

Saunders, C. T., et al. (2012). "Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs." Bioinformatics **28**(14): 1811-1817.

Sommer, A., et al. (2006). "Clinical and serological characteristics of progressive facial hemiatrophy: a case series of 12 patients." Journal of the American Academy of Dermatology **54**(2): 227-233.

SourceForge "Burrows-Wheeler Alignment Tool."

Stone, J. (2003). "Parry–Romberg syndrome A global survey of 205 patients using the Internet." Neurology **61**(5): 674-676.

Szolek, A., et al. (2014). "OptiType: precision HLA typing from next-generation sequencing data." Bioinformatics **30**(23): 3310-3316.

Szramka-Pawlak, B., et al. (2014). "Quality of life and optimism in patients with morphea." Applied research in quality of life **9**(4): 863-870.

Thapar, A. and M. Cooper (2013). "Copy number variation: what is it and what has it told us about child psychiatric disorders?" Journal of the American Academy of Child & Adolescent Psychiatry **52**(8): 772-774.

Thavarajah, R., et al. (2012). "Chemical and physical basics of routine formaldehyde fixation." Journal of oral and maxillofacial pathology: JOMFP **16**(3): 400.

Thorvaldsdóttir, H., et al. (2013). "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration." Briefings in bioinformatics **14**(2): 178-192.

Tolkachjov, S. N., et al. (2015). "Progressive hemifacial atrophy: a review." Orphanet journal of rare diseases **10**(1): 39.

Tollefson, M. M. and P. M. Witman (2007). "En coup de sabre morphea and Parry-Romberg syndrome: a retrospective review of 54 patients." Journal of the American Academy of Dermatology **56**(2): 257-263.

Torok, K. S. and T. Arkachaisri (2012). "Methotrexate and corticosteroids in the treatment of localized scleroderma: a standardized prospective longitudinal single-center study." The Journal of rheumatology **39**(2): 286-294.

Voelkerding, K. V., et al. (2009). "Next-generation sequencing: from basic research to diagnostics." Clinical chemistry **55**(4): 641-658.

Wada, T., et al. (2003). "A Japanese case of segmental Darier's disease caused by mosaicism for the ATP2A2 mutation." British Journal of Dermatology **149**(1): 185-188.

Wang, B., et al. (2010). "γ-secretase gene mutations in familial acne inversa." Science **330**(6007): 1065-1065.

Wang, K., et al. (2010). "ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data." Nucleic acids research **38**(16): e164-e164.

Weibel, L. and J. Harper (2008). "Linear morphoea follows Blaschko's lines." British Journal of Dermatology **159**(1): 175-181.

Weibel, L., et al. (2011). "Misdiagnosis and delay in referral of children with localized scleroderma." British Journal of Dermatology **165**(6): 1308-1313.

Weibel, L., et al. (2006). "Evaluation of methotrexate and corticosteroids for the treatment of localized scleroderma (morphoea) in children." British Journal of Dermatology **155**(5): 1013-1020.

Wong, S. Q., et al. (2014). "Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing." BMC medical genomics **7**(1): 23.

Wu, M. C., et al. (2011). "Rare-variant association testing for sequencing data with the sequence kernel association test." The American Journal of Human Genetics **89**(1): 82-93.

Yamanaka, C. and N. Gibbs (1999). "Trauma-induced linear scleroderma." Cutis **63**(1): 29-32.

Zulian, F., et al. (2005). "Juvenile localized scleroderma: clinical and epidemiological features in 750 children. An international study." Rheumatology **45**(5): 614-620.

Zulian, F., et al. (2011). "Methotrexate treatment in juvenile localized scleroderma: A randomized, double-blind, placebo-controlled trial." Arthritis & Rheumatology **63**(7): 1998-2006.

Zulian, F., et al. (2006). "Congenital localized scleroderma." The Journal of pediatrics **149**(2): 248-251.

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Dr. med. Alexander Navarini. Not just for the chance to do this work but also for taking the time to go above and beyond to teach skills and advise me in many matters. I have learned so much from him being my mentor these last four years and I have enjoyed every minute of it. I sincerely hope this is not the last time we work together.

Thank you to my committee members, Prof. Dr. Burkhard Becher and Prof. Dr. med. Anita Rauch. Their support and input in these last four years has helped the success of the project and taught me a lot.

Thank you to the clinical trial doctors who collected the samples and the clinical data that made this project possible, Dr. Regula Wälchli, Dr. Martin Theiler and Dr. Alexandra Smith. Above all I am grateful to Dr. Lisa Weibel whose preceding work on LLS paved the way for this project and who advised throughout.

There are several friends in dermatology I need to thank. First and foremost, Phil Cheng who was always there to lend an ear and frequently helped me with bioinformatics analysis. Several of my breakthroughs were a direct result from our conversations and he only ever accepted payments in cups of coffee. Second, Adhideb Ghosh, who selflessly helped with bioinformatics when I asked for it no matter how busy he was. Thank you to my friends Sabrina, Alshimaa and Ishani for support and laughs throughout the years.

I am indebted to the many professors and collaborators who have advised me along the way. Prof. Dr. Mike Recher, Prof. Dr. Christoph Hess, PD. Dr. med. Emmanuella Guenova, Prof. Dr. Lukas Flatz, Dr. Andrew Pink and Prof. Dr. med. Nikhil Yawalkar my collaborators who allowed

me to contribute to projects outside of my field allowing me to learn about different diseases and expand my mind. Prof. Dr. French, Prof. Dr. Dummer, Prof. Dr. Jürg Hafner and Prof. Dr. Levesque, all of whom contributed and advised my project throughout the years at my presentations.

I owe a huge thank you to my family. To my parents, Sheila and Fergus who always took a phone call from me to share my achievements or lift me up in my low moments. They never hesitated to support me in all my educational adventures from space camp to a master degree to starting a new life in Zürich. To my sister, Aislinn who works tirelessly to make sure I am happy and never needs to be prompted to support me. I eventually managed to undo the damage she did to my understanding of the metric system and go on to be a scientist. And to Gavin, my brother-in-law who without fail can always make me laugh.

A special thank you to my partner, Marco, who has made sure I am happy and healthy when I am too distracted or busy to do so myself. He always makes sure that no matter how hectic the day, I always go to sleep smiling and feeling loved.

To my friends abroad, Tonya, Grace and Orla with whom I have shared my most important moments since our bachelor. Rachel, Aoife and Ruth, my oldest friends. Though none of you live close to me anymore you are a reservoir of love, advice and laughs. Thanks to my friends in Zürich, Francesca, Wilson, Jilllianne, Corina and Daniel amongst others for being my break from PhD life with travel and friendship. Thank you to my friends Ossia and Nina for translating my abstract to German and Melanie for helping me in the lab.

Go raibh maith agaibh go léir le haghaidh do chabhair agus cairdeas na ceithre bliana seo.

Curriculum Vitae

HIGGINS, Rebecca Geraldine

Date of Birth: 06/12/1989

Personal Phone: +41 787380870

Email: rebecca.higgins@usz.ch

Home town: Athlone, Westmeath

Canton: Zürich

Nationality: Irish

EDUCATION

September 2002 - May 2008: Secondary School. Our Lady's Bower, Athlone, Westmeath, Ireland.

Leaving Certificate Results (2008): Points: 485. Biology, Chemistry, Maths, Accounting, French, Irish and English.

September 2008 - June 2012: Undergraduate in Natural Sciences (BA mod. Microbiology). Trinity College Dublin, Dublin, D2, Ireland.
Graduation date: 01/11/2012

- Year 1 (Sept. 2008 - May 2009) School of Natural Sciences. Subjects included Biology, Maths Methods, Chemistry, Foundation Physics
- Year 2 (Sept. 2009 - May 2010) Faculty of Science. Chemistry I and II including the organic, inorganic and physical chemistries, Biology I and II (including Cell Structure and Function, Metabolism, Genetics, Microbiology, Physiology, Animal Behaviour, Infection and Immunity)
- Year 3 and 4 (Sept. 2010 - June 2012) Department of Microbiology. Moderatorship in Microbiology. Subjects include Genetics, Biochemistry, Microbiology, Cellular Biology and Molecular Biology modules.

September 2012 - August 2013: Master Degree (MSc). University College Dublin, Belfield, Dublin 4, Ireland. Graduation date: 11/2013

Masters with a focus on Evolutionary Biology. School of Biology and Environmental Science.

Project title: "The Biogeographical Origins of *Ilex aquifolium*". (Supervisor: Dr. Colin Kelleher.) Thesis based in the National Botanic Gardens, Dublin.

April 2014 - present

Employed as a PhD student at UZH since 01/04/2014**.** University of Zurich, Rämistrasse 74, 8006, Zurich.

Life Science Zurich Graduate School UZH; Microbiology and Immunology Program.

Supervisor: Prof. Dr. med. Dr. sc. nat. Alexander Navarini, University Hospital Zurich, Department of Dermatology. alexander.navarini@usz.ch

PUBLICATIONS

Higgins R, et al. "Generalized Comedones, acne, and Hidradenitis suppurativa in a patient with an FGFR2 Missense Mutation." Frontiers in Medicine 4 (2017).

Navarini AA., Hruz P, Berger CT, Hou TZ, Schwab C, Gabrysch A, **Higgins R**, Frede N, Sgier BCP, Kämpe O, Burgener AV. 2016. Vedolizumab as a successful treatment of CTLA-4–associated autoimmune enterocolitis. *Journal of Allergy and Clinical Immunology*.

Marquardsen F, Baldin F, Wunderer F, Herz WA, Mikhael R, Baz Z, Rezaee F, Hanna F,Meyer B, Jauch A, Bigler M, Burgener AV, **Higgins R,** Navarini AA, Geha R, Notarangelo1 LD, Hess C, Berger CT, Bloch DB, Recher M. “Detection of Sp110 by Flow Cytometry and Application to Screening Patients for Veno-occlusive Disease with Immunodeficiency." Journal of clinical immunology 37.7 (2017): 707-714.